

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number  
**WO 01/25438 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 15/21, C07K 14/56, C12N 1/21, A61K 31/70, C12N 15/62, C07K 16/24, A61K 38/21, C12N 5/10, G06F 19/00**

(74) Agents: QUINE, Jonathan, Alan; The Law Offices of Jonathan Alan Quine, P.O. Box 458, Alameda, CA 94501 et al. (US).

(21) International Application Number: PCT/US00/27781

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 6 October 2000 (06.10.2000)

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English

Published:

— Without international search report and to be republished upon receipt of that report.

(26) Publication Language: English

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(30) Priority Data:  
09/415,183 7 October 1999 (07.10.1999) US

(71) Applicant (*for all designated States except US*): MAXY-GEN, INC. [US/US]; 515 Galveston Drive, Redwood City, CA 94063 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): HEINRICH, Volker [DE/US]; Apartment 9, 1915 Mount Vernon Court, Mountain View, CA 94040 (US). CHEN, Teddy [US/US]; 3652 McNulty Way, Redwood City, CA (US). PATTEN, Phillip, A. [US/US]; 261 La Cuesta Drive, Menlo Park, CA 94025 (US).



**WO 01/25438 A2**

(54) Title: IFN-ALPHA HOMOLOGUES

(57) Abstract: Alpha interferon homologues (both nucleic acids and polypeptides) are provided. Compositions including these interferon homologue polypeptides and nucleic acids, recombinant cells comprising said homologue polypeptides and nucleic acids, methods of making the new homologues, antibodies to the new homologues, and methods of using the homologues are provided. Integrated systems comprising the sequences of the nucleic acids or polypeptides are also provided.

**IFN-ALPHA HOMOLOGUES****COPYRIGHT NOTIFICATION**

Pursuant to 37 C.F.R. 1.71(e), a portion of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

10

**CROSS REFERENCE TO RELATED APPLICATIONS**

This application is a continuation-in-part application of and claims the benefit of and priority to U.S. Patent Application Serial No. 09/145,483, filed October 7, 1999, the disclosure of which is incorporated herein by reference in its entirety for all purposes.

15

**FIELD OF THE INVENTION**

The present invention relates to the generation of new interferon-alpha homologues.

20

**BACKGROUND OF THE INVENTION**

Interferon-alphas are members of the diverse helical-bundle superfamily of cytokine genes (Sprang, S.R. *et al.* (1993) *Curr. Opin. Struct. Biol.* 3:815-827). The human interferon-alphas are encoded by a family of over 20 tandemly duplicated nonallelic genes that share 85-98% sequence identity at the amino acid level (Henco, K. *et al.* (1985) *J. Mol. Biol.* 185:227-260).

25

Interferon-alphas have been shown to inhibit various types of cellular proliferation, and are especially useful for the treatment of a variety of cellular proliferation disorders frequently associated with cancer, particularly hematologic malignancies such as leukemias. These proteins have shown antiproliferative activity against multiple myeloma, chronic lymphocytic leukemia, low-grade lymphoma, Kaposi's sarcoma, chronic myelogenous leukemia, renal-cell carcinoma, urinary bladder tumors and ovarian cancers (Bonnam, E.M. *et al.* (1984) *J. Biol. Response Modifiers* 3:580; Oldham, R.K. (1985) *Hospital Practice* 20:71).

30

Interferon-alphas are also useful against various types of viral infections (Finter, N.B. *et al.* (1991) *Drugs* 42(5):749). Interferon-alphas have shown activity against human papillomavirus infection, Hepatitis B, and Hepatitis C infections (Finter, N.B. *et al.*, 1991, *supra*; Kashima, H. *et al.* (1988) *Laryngoscope* 98:334; Dusheiko, G.M. 5 *et al.* (1986) *J. Hematology* 3 (Supple. 2):S199; Davis, GL *et al.* (1989) *N. England J. Med.* 321:1501). The role of interferons and interferon receptors in the pathogenesis of certain autoimmune and inflammatory diseases has also been investigated (Benoit, P. *et al.* (1993) *J. Immunol.* 150(3):707).

Although these proteins possess therapeutic value in the treatment of a 10 number of diseases, they have not been optimized for use as pharmaceuticals. For example, dose-limiting toxicity, receptor cross-reactivity, and short serum half-lives significantly reduce the clinical utility of many of these cytokines (Dusheiko, G. (1997) *Hepatology* 26:112S–121S; Vial, T. and Descotes, J. (1994) *Drug Experience* 10:115–150; Funke, I. *et al.* (1994) *Ann. Hematol.* 68:49–52; Schomburg, A. *et al.* (1993) *J. 15 Cancer Res. Clin. Oncol.* 119:745–755). Diverse and severe side effect profiles which accompany interferon administration include flu-like symptoms, fatigue, neurological disorders including hallucination, fever, hepatic enzyme elevation, and leukopenia (Pontzer, C.H. *et al.* (1991) *Cancer Res.* 51:5304; Oldham, 1985, *supra*).

The existence of abundant naturally occurring sequence diversity within the 20 interferon-alphas (and hence a large sequence space of recombinants) along with the intricacy of interferon-alpha/receptor interactions and variety of therapeutic and prophylactic activities creates an opportunity for the construction of superior interferon homologues.

#### SUMMARY OF THE INVENTION

25 The invention provides novel interferon-alpha (IFN-alpha or IFN- $\alpha$ ) homologue polypeptides, nucleic acids encoding the polypeptides and complementary nucleotide sequences thereof, fragments of said polypeptides and nucleic acids, antibodies to the polypeptides, and uses therefor, data sets containing character strings of interferon-alpha homologue sequences, and automated systems for using the character strings.

30 In one aspect, the invention includes an isolated or recombinant interferon-alpha nucleic acid homologue. Included are a polynucleotide sequences selected from SEQ ID NO:1 to SEQ ID NO:35, or to SEQ ID NO:72 to SEQ ID NO:78, and

complementary polynucleotide sequences thereof. Polynucleotide sequences encoding a polypeptide selected from SEQ ID NO:36 to SEQ ID NO:81 or from SEQ ID NO:79 to SEQ ID NO:85, and complementary polynucleotide sequences thereof are also a feature of the invention. Similarly, a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of any of the preceding polynucleotide sequences is a feature of the present invention. In addition, a polynucleotide sequence comprising a nucleotide fragment of any of the preceding polynucleotide sequences which nucleotide fragment encodes a polypeptide having an antiproliferative activity in a human Daudi cell line- based cell proliferation assay is a feature of the invention. Similarly, a polynucleotide sequence comprising a nucleotide fragment of any of the polynucleotide sequences of the invention described above and below which encodes a polypeptide having antiviral activity in a murine cell line/EMCV - based assay is a feature of the invention.

The invention also includes an isolated or recombinant nucleic acid, comprising a polynucleotide sequence encoding a polypeptide, wherein the polypeptide comprises the amino acid sequence: CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QQLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof, where X<sub>11</sub> is N or D; X<sub>12</sub> is R, S, or K; X<sub>15</sub> is L or M; X<sub>16</sub> is I, M, or V; X<sub>19</sub> is A or G; X<sub>22</sub> is G or R; X<sub>24</sub> is I or T; X<sub>26</sub> is P or H; X<sub>34</sub> is H, Y or Q; X<sub>38</sub> is F or L; X<sub>40</sub> is Q or R; X<sub>45</sub> is G or S; X<sub>46</sub> is N or H; X<sub>47</sub> is Q or R; X<sub>50</sub> is K or R; X<sub>51</sub> is A or T; X<sub>55</sub> is S or F; X<sub>56</sub> is V or A; X<sub>57</sub> is L or F; X<sub>60</sub> is M or I; X<sub>61</sub> is I or M; X<sub>67</sub> is L or F; X<sub>72</sub> is D or N; X<sub>75</sub> is A or V; X<sub>76</sub> is A or T; X<sub>78</sub> is E or D; X<sub>79</sub> is Q or E; X<sub>80</sub> is S, R, T, or N; X<sub>83</sub> is E or D; X<sub>85</sub> is F or L; X<sub>86</sub> is S or Y; X<sub>88</sub> is E or G; X<sub>90</sub> is Y, H, N; X<sub>95</sub> is D, E, or N; X<sub>101</sub> is I, M, or V; X<sub>103</sub> is E or G; X<sub>105</sub> is G or W; X<sub>106</sub> is V or M; X<sub>107</sub> is E, G, or K; X<sub>108</sub> is E or G; X<sub>114</sub> is V, E, or G; X<sub>116</sub> is S or P; X<sub>121</sub> is K or R; X<sub>124</sub> is F or L; X<sub>132</sub> is T, I, or M; X<sub>134</sub> is K or R; and X<sub>140</sub> is A or S. Each of the single letters of this amino acid sequence represents a particular amino acid residue according to standard practice known to those of ordinary skill in the art.

A polypeptide having any of the preceding sequences, such as those embodied in SEQ ID NO:36 to SEQ ID NO:54, is also a feature of the invention.

In other embodiments, the encoded polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:36 to SEQ ID NO:54; and the  
5 nucleic acid comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 to SEQ ID NO:19.

The invention also provides polypeptide fragments of any of SEQ NOS:36-  
70 and SEQ ID NOS:72-79. In one aspect of the invention, such a polypeptide fragment exhibits an antiproliferative activity in a human Daudi cell line- based cell proliferation  
10 assay or an antiviral activity in a murine cell line/EMCV - based assay, or both said activities. The human Daudi cell line- based cell proliferation assay and antiviral activity in a murine cell line/EMCV - based assay are described in greater detail below. In yet another aspect, the invention provides a polynucleotide sequence comprising a nucleotide fragment of any nucleic acid of the invention described above and below, wherein said  
15 nucleotide fragment encodes a polypeptide fragment that exhibits an antiproliferative activity in a human Daudi cell line- based cell proliferation assay or an antiviral activity in a murine cell line/EMCV - based assay, or both activities, as is described in greater detail below.

The invention also includes an isolated or recombinant nucleic acid  
20 comprising a polynucleotide sequence encoding a polypeptide, wherein the polypeptide comprises an amino acid sequence comprising at least 20 contiguous amino acids of any one of SEQ ID NOS:36-70. In other embodiments, the polypeptide of the invention comprises an amino acid sequence comprising one or more of amino acid residues (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and  
25 Glu166, wherein the numbering of the amino acid residues corresponds to the numbering of residues in the amino acid sequence of SEQ ID NO:36. In various embodiments, the encoded polypeptide of the invention comprises at least 30, at least 50, at least 70, at least 75, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 155, at least 160, or at least 165 contiguous amino acid residues of any one of SEQ ID  
30 NOS:36-70. In other embodiments, the encoded polypeptide is at least 150, at least 155, at least 160, at least 163, or at least 165 amino acids in length. In another embodiment, the encoded polypeptide is about 166 amino acids in length. In yet other embodiments, the

encoded polypeptide comprises an amino acid sequence selected from SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:45, and SEQ ID NO:46.

In other embodiments, the invention provides a nucleic acid that comprises  
5 a polynucleotide sequence selected from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:10, and SEQ ID NO:11.

In other embodiments, the polypeptide encoded by any nucleic acid or the invention described herein or a fragment thereof may have antiproliferative activity in a  
10 human Daudi cell line - based assay, or antiviral activity in a human WISH cell/EMCV-based assay. In other embodiments, the encoded polypeptide has antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in the human Daudi cell line - based assay (1 unit is the amount of protein in milligram (mg) required to induce 50% antiproliferative activity), or antiviral activity of at about least  $2.1 \times 10^7$  units/milligram (mg) in the human  
15 WISH cell/EMCV-based assay (1 unit is the amount of protein in mg required to induce 50% antiviral activity). In other embodiments, the encoded polypeptide can bind to a type I interferon receptor, preferably a human type I interferon receptor, more preferably a human (*e.g.*, type I) interferon-alpha receptor.

The invention also includes a cell comprising any nucleic acid of the  
20 invention described herein, or which expresses any polypeptide of the invention noted herein. In one embodiment, the cell expresses a polypeptide encoded by the nucleic acid of the invention as described herein.

The invention also includes a vector comprising any nucleic acid of the invention described above and below. The vector can comprise a plasmid, a cosmid, a  
25 phage, or a virus; the vector can be, *e.g.*, an expression vector, a cloning vector, a packaging vector, an integration vector, or the like. The invention also includes a cell transduced by a vector of the invention. The invention also includes compositions comprising any nucleic acid of the invention described above and below, and an excipient, preferably a pharmaceutically acceptable excipient. Cells and transgenic animals which  
30 include any polypeptide or nucleic acid of the invention described above and below, *e.g.*, produced by transduction of vector, are a feature of the invention.

The invention also includes compositions produced by digesting one or more of the nucleic acids of the invention described above or below with a restriction endonuclease, an RNase, or a DNase; and, compositions produced by incubating one or more nucleic acids described above or below in the presence of deoxyribonucleotide triphosphates and a nucleic acid polymerase, *e.g.*, a thermostable polymerase.

5 The invention also includes compositions comprising two or more nucleic acids described above or below. The composition may comprise a library of nucleic acids, where the library contains at least about 5, 10, 20 or 50 nucleic acids.

In another aspect, the invention includes an isolated or recombinant 10 polypeptide encoded by any nucleic acid described above or below. In one embodiment, the polypeptide may comprise a sequence selected from SEQ ID NO:36 to SEQ ID NO:70, or SEQ ID NO:79 to SEQ ID NO:85.

The invention also includes a polypeptide comprising at least 50 contiguous 15 amino acids of a protein encoded by a polynucleotide sequence, the polynucleotide sequence selected from the group consisting of: (a) SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78; (b) a polynucleotide sequence that encodes a polypeptide selected from SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85; and (c) a complementary sequence of a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide 20 sequence (a) or (b). In various embodiments, the polypeptide comprises at least about 70, 100, 120, 130, 140, 150, 155, 160, 165, or 166 contiguous amino acids of the encoded protein.

The invention also includes an isolated or recombinant polypeptide comprising an amino acid sequence comprising at least 50 contiguous amino acid residues 25 of any one of SEQ ID NOS:36-70, and one or more of amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, where the numbering of the amino acids corresponds to that of SEQ ID NO:36. In various 30 embodiments, the polypeptide comprises at least about 50, 70, 75, 100, 110, 120, 130, 140, 150, 155, 160, 163, 165, or 166 contiguous amino acids of any one of SEQ ID NOS:36-70. In more preferred embodiments, the polypeptide comprises at least about 50, 70, 75, 100, 110, 120, 130, 140, 150, 155, 160, 163, 165, or 166 contiguous amino acid residues of any one of SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41,

SEQ ID NO:42, SEQ ID NO:45, or SEQ ID NO:46. In other embodiments, the polypeptide of the invention is at least about 50, 70, 75, 100, 110, 120, 130, 140, 150, 155, 160, 163, 165, or 166 amino acid residues in length, or is preferably 166 amino acids in length. Longer polypeptides, *e.g.*, which comprise purification tags or the like, are also 5 contemplated. Such polypeptides may display antiproliferative activities in human Daudi cell-line based assay and/or antiviral activities in a human WISH cell/EMCV-based assay.

The invention also includes a polypeptide which specifically binds polyclonal antisera raised against at least one antigen, said at least one antigen comprising a polypeptide sequence selected from an amino acid sequence set forth in SEQ ID NO:36 10 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85 or a fragment thereof. In particular, the invention provides polypeptides which bind a polyclonal antisera raised against at least one antigen, wherein said at least one antigen comprises at least one amino acid sequence set forth in SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, or a fragment of any of these amino sequences, wherein the polyclonal antisera is 15 subtracted with one or more known interferon-alpha polypeptides or proteins, including, *e.g.*, a polypeptide or protein encoded by a nucleic acid having or corresponding to one or more of the following GenBank™ accession numbers: J00210 (alpha-D), J00207 (Alpha-a), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 20 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), and other similar or homologous interferon-alpha nucleic acid sequences presented in GenBank.

Any polypeptide described above or below optionally has antiproliferative 25 activity in a human Daudi cell line - based assay and/or in an antiviral activity in a human WISH cell/EMCV-based assay. Any polypeptide described above or below can have antiproliferative activity of at least about  $8.3 \times 10^6$  units/mg in the human Daudi cell line - based assay or antiviral activity of at least about  $2.1 \times 10^7$  units/mg in the human WISH cell/EMCV-based assay. In other embodiments, any polypeptide described above or 30 below can bind to a type I interferon receptor, preferably a human type I interferon receptor, more preferably a human interferon-alpha receptor.

In other embodiments, any polypeptide described above or below may further include a secretion/localization sequence, *e.g.*, a signal sequence, an organelle targeting sequence, a membrane localization sequence, and the like. Any polypeptide described herein may further include a sequence that facilitates purification, *e.g.*, an epitope tag (such as, a FLAG epitope), a polyhistidine tag, a GST fusion, and the like.

5 The polypeptide optionally includes a methionine at the N-terminus. Any polypeptide of the invention described herein optionally includes one or more modified amino acids, such as a glycosylated amino acid, a PEG-ylated amino acid, a farnesylated amino acid, an acetylated amino acid, a biotinylated amino acid, a carboxylated amino acid, a

10 phosphorylated amino acid, an acylated amino acid, or the like.

The invention also includes compositions comprising any polypeptide described herein in an excipient, preferably a pharmaceutically acceptable excipient.

The invention also includes an antibody or antisera produced by administering one or more of the polypeptides of the invention described herein to a mammal, wherein the antibody or antisera does not specifically bind to a known alpha-interferon polypeptide or protein, including, *e.g.*, any polypeptide or protein encoded by a nucleic acid having or corresponding to one or more of the following GenBank accession numbers: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), and other similar or homologous interferon-alpha sequences presented in GenBank.

The invention also includes antibodies which specifically bind a polypeptide comprising a sequence selected from SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85. The antibodies are, *e.g.*, polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments, fragments produced by an Fab expression library, or the like.

Methods for producing the polypeptides of the invention are also included.

30 One such method comprises introducing into a population of cells any nucleic acid described herein, operatively linked to a regulatory sequence effective to produce the encoded polypeptide, culturing the cells in a culture medium to produce the polypeptide,

and optionally isolating the polypeptide from the cells or from the culture medium. The nucleic acid may be part of a vector, such as a recombinant expression vector.

The invention also includes a method of inhibiting growth of tumor cells, by contacting the tumor cells with a polypeptide of the invention described herein, thereby inhibiting growth of the tumor cells. In one embodiment, the invention includes a method of inhibiting growth of population of tumor cells comprising contacting the population of tumor cells with an effective amount of a polypeptide of the invention sufficient to inhibit growth of tumor cells in said population of tumor cells, thereby inhibiting growth of tumor cells in said population of cells. In various embodiments, the tumor cells can be human carcinoma cells, human leukemia cells, human T-lymphoma cells, human melanoma cells, other human cancer cells as described herein, and the like. The tumor cells can be *in vivo*, *ex vivo*, or *in vitro* (e.g., cultured cells).

The invention also includes a method of inhibiting the replication of a virus within one or more cells infected by the virus, by contacting one or more of the infected cells with an effective amount of a polypeptide of the invention as described above and below, wherein said amount is sufficient to inhibit viral replication in said one or more infected cells, thereby inhibiting replication of the virus in the one or more cells. In various embodiments, the virus can be an RNA virus, e.g., a human immunodeficiency virus or a hepatitis C virus, or a DNA virus, e.g., a hepatitis B virus. The infected cells can be *in vivo*, *ex vivo*, or *in vitro* (e.g., cultured cells).

The invention also includes a method of treating an autoimmune disorder in a subject in need of such treatment, by administering to the subject an effective amount of a polypeptide of the invention as described herein sufficient to treat the autoimmune disorder. In various embodiments, the autoimmune disorder may be multiple sclerosis, rheumatoid arthritis, lupus erythematosus, type I diabetes, and the like. The invention also includes, in a method of treating a disorder treatable by administration of interferon-alpha to a subject, an improvement comprising administering to the subject an effective amount of a polypeptide of the invention as described herein sufficient to treat said disorder. The disorder treatable by administration of interferon-alpha disorder may be multiple sclerosis, rheumatoid arthritis, lupus erythematosus, type I diabetes, AIDS or AIDS-related complexes, or the like.

In general, nucleic acids and proteins derived by mutation of the sequences herein are a feature of the invention. Similarly, those produced by diversity generation or recursive sequence recombination (RSR) methods (*e.g.*, DNA shuffling) are a feature of the invention. Mutation and recombination methods using the nucleic acids described 5 herein are a feature of the invention. For example, one method of the invention includes recursively recombining one or more nucleic acid sequences of the invention as described above and below with one or more additional nucleic acids (including, but not limited to, those noted herein), each sequence of the one or more additional nucleic acids encoding an interferon-alpha homologue or an amino acid subsequence thereof. The recombining steps 10 are optionally performed *in vivo*, *ex vivo*, *in silico* or *in vitro*. Said recursive recombination produces at least one library of recombinant interferon-alpha homologue nucleic acids. Also included in the invention are a recombinant interferon-alpha homologue nucleic acid produced by this method, a cell containing the recombinant interferon-alpha homologue nucleic acid, a nucleic acid library produced by this recursive 15 recombination method, a composition comprising two or more of said recombinant interferon-alpha nucleic acids, and a population of cells comprising such recombinant interferon-alpha nucleic acids or containing the library. In one embodiment, the library comprise at least ten such recombinant nucleic acids.

The invention also provides a method of producing a modified or 20 recombinant interferon-alpha homologue nucleic acid that comprises mutating a nucleic acid of the invention as described herein.

Also provided are nucleic acids that encode an interferon-alpha homologue having an increased growth inhibition activity, cytostatic activity, or cytotoxic activity against a population of cells (*e.g.*, cancer cells) relative to the growth inhibition activity 25 cytostatic activity, or cytotoxic activity, respectively, of human interferon-alpha 2a or other known interferon-alpha against the population of cells.

These and other objects and features of the invention will become more fully apparent when the following detailed description is read in conjunction with the accompanying figures.

**BRIEF DESCRIPTION OF THE FIGURES**

Figures 1A-1E show an alignment of exemplary mature interferon homologue polypeptide sequences (SEQ ID NOS: 36-70 and 79-85) according to the invention.

5       Figure 2 shows antiproliferative activities in a human Daudi cell line - based assay and antiviral activities in a human WISH cell/EMCV-based assay of, respectively, exemplary interferon homologues of the present invention relative to the respective antiproliferative and antiviral activities of two control compounds, human interferon alpha-2a ("IFN- $\alpha$ -2a" or "2a") and consensus human interferon ("IFN-Con1" or 10 "Con1").

Figures 3A, 3B, and 3C illustrate activity profiles of IFN-alpha homologue 3DA11 (SEQ ID NO:40) and control interferons, human interferon alpha-2a ("2a") and consensus human interferon alpha ("Con1"), against a panel of tumor cell lines. Fig. 3A shows the cell total growth inhibitory activity of IFN-alpha homologue 3DA11 and each 15 control IFN on each respective cell line as reflected in the GI50 value, which is the concentration ( $\mu\text{g/ml}$ ) of interferon alpha homologue or control IFN alpha at which growth of a particular cell line is inhibited by 50%, as measured by a 50% reduction in the net protein/polypeptide increase in the interferon alpha homologue or control IFN alpha at the end of the incubation period.

20       Fig. 3B shows the cytostatic activity of IFN-alpha homologue 3DA11 and each control IFN on each cell line of the panel of cell lines. Cytostatic activity refers to an activity capable of suppressing growth and multiplication of cells. Cytostatic activity is assessed as a reflection of the concentration of IFN-alpha homologue 3DA11 or control IFN ( $\mu\text{g/ml}$ ) at which the growth and/or multiplication of cells of a particular cell line is 25 completely inhibited or suppressed, such that the amount of cellular protein at the end of the incubation period equals the amount of cellular protein at the beginning of the incubation period ("total growth inhibition" or "TGI").

Fig. 3C illustrates the cytotoxic activity of IFN-alpha homologue 3DA11 and each control IFN on each respective cell line. The cytotoxicity of an agent (e.g., an 30 IFN homologue or IFN compound) is the degree to which the agent possess a specific destructive action on certain cells or the possession of such action. The term typically refers to an agent capable of causing cell death and is used particularly in referring to the lysis of cells by immune phenomena and to agents of compounds that selectively kill

dividing cells. In Fig. 3C, cytotoxic activity is illustrated as LC50, the concentration of IFN-alpha homologue 3DA11 ( $\mu\text{g}/\text{ml}$ ) at which a 50% reduction in the net protein increase in control cells (control IFN alpha) at the end of the incubation as compared to that at the beginning of the incubation period is observed, indicating a net loss of cells following  
5 addition of the particular interferon. Cytotoxic activity may be assessed as the concentration of IFN-alpha homologue 3DA11 at which, relative to the control cells, 50% of the total number of cells (*i.e.*, total population) of a particular cell line are destroyed or killed.

Figs. 4A, 4B, 4C, and 4D show the cytostatic activity of selected  
10 interferon-alpha homologues of the present invention relative to the cytostatic activities of two control interferon alphas, human interferon-alpha 2a ("2a") and consensus human interferon-alpha ("Con1"), against a leukemia cell line (RPMI-8226) (Fig. 4A), a lung cancer cell line (NCI-H23) (Fig. 4B), a renal cancer cell line (ACHN) (Fig. 4C), and an ovarian cancer cell line (OVCAR-3) (Fig. 4D), respectively. Cytostatic activity is  
15 reflected by a TGI value for a particular interferon alpha (*i.e.*, the concentration of interferon alpha at which cell growth of a cell line is totally inhibited, wherein the amount of cellular protein at the end of the incubation period equals the amount of cellular protein at the beginning of the incubation period).

Fig. 5 presents a comparison of the number of mice (out of a total number  
20 of six mice) that survived following administration of doses of 2  $\mu\text{g}$ , 10  $\mu\text{g}$ , and 50  $\mu\text{g}$  of two exemplary IFN-alpha homologues of the present invention (designated "IFN-CH2.2" and "IFN-CH2.3"), doses of 2  $\mu\text{g}$ , 10  $\mu\text{g}$ , and 50  $\mu\text{g}$  of murine IFN-alpha-4, and doses of 2  $\mu\text{g}$ , 10  $\mu\text{g}$ , and 50  $\mu\text{g}$  of human IFN-alpha-2a, respectively. The results shown in Fig. 5 demonstrate that in a murine model system, the improved *in vitro* antiviral activity of these  
25 two exemplary IFN-alpha homologues is maintained and sustained *in vivo*. Phosphate-buffered saline (PBS) is used as a control.

## DETAILED DESCRIPTION OF THE INVENTION

### DEFINITIONS

30 Unless otherwise defined herein or below in the remainder of the specification, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the present invention belongs.

A "polynucleotide sequence" is a nucleic acid (which is a polymer of nucleotides (A,C,T,U,G, etc. or naturally occurring or artificial nucleotide analogues)) or a character string representing a nucleic acid, depending on context. Either the given nucleic acid or the complementary nucleic acid can be determined from any specified 5 polynucleotide sequence.

Similarly, an "amino acid sequence" is a polymer of amino acids (a protein, polypeptide, *etc.*) or a character string representing an amino acid polymer, depending on context. Either the given nucleic acid or the complementary nucleic acid can be determined from any specified polynucleotide sequence.

10 A nucleic acid, protein, peptide, polypeptide, or other component is "isolated" when it is partially or completely separated from components with which it is normally associated (other peptides, polypeptides, proteins (including complexes, *e.g.*, polymerases and ribosomes which may accompany a native sequence), nucleic acids, cells, synthetic reagents, cellular contaminants, cellular components, *etc.*), *e.g.*, such as 15 from other components with which it is normally associated in the cell from which it was originally derived. A nucleic acid, polypeptide, or other component is isolated when it is partially or completely recovered or separated from other components of its natural environment such that it is the predominant species present in a composition, mixture, or collection of components (*i.e.*, on a molar basis it is more abundant than any other 20 individual species in the composition). In preferred embodiments, the preparation consists of more than 70%, typically more than 80%, or preferably more than 90% of the isolated species.

In one aspect, a "substantially pure" or "isolated" nucleic acid (*e.g.*, RNA or DNA), polypeptide, protein, or composition also means where the object species (*e.g.*, 25 nucleic acid or polypeptide) comprises at least about 50, 60, or 70 percent by weight (on a molar basis) of all macromolecular species present. A substantially pure or isolated composition can also comprise at least about 80, 90, or 95 percent by weight of all macromolecular species present in the composition. An isolated object species can also be purified to essential homogeneity (contaminant species cannot be detected in the 30 composition by conventional detection methods) wherein the composition consists essentially of derivatives of a single macromolecular species.

The term "isolated nucleic acid" may refer to a nucleic acid (*e.g.*, DNA or RNA) that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (*i.e.*, one at the 5' and one at the 3' end) in the naturally occurring genome of the organism from which the nucleic acid of the invention is derived.

- 5 Thus, this term includes, *e.g.*, a cDNA or a genomic DNA fragment produced by polymerase chain reaction (PCR) or restriction endonuclease treatment, whether such cDNA or genomic DNA fragment is incorporated into a vector, integrated into the genome of the same or a different species than the organism, including, *e.g.*, a virus, from which it was originally derived, linked to an additional coding sequence to form a hybrid gene  
10 encoding a chimeric polypeptide, or independent of any other DNA sequences. The DNA may be double-stranded or single-stranded, sense or antisense.

A nucleic acid or polypeptide is "recombinant" when it is artificial or engineered, or derived from an artificial or engineered protein or nucleic acid. The term "recombinant" when used with reference *e.g.*, to a cell, nucleotide, vector, or polypeptide  
15 typically indicates that the cell, nucleotide, or vector has been modified by the introduction of a heterologous (or foreign) nucleic acid or the alteration of a native nucleic acid, or that the polypeptide has been modified by the introduction of a heterologous amino acid, or that the cell is derived from a cell so modified. Recombinant cells express nucleic acid sequences (*e.g.*, genes) that are not found in the native (non-recombinant) form of the cell  
20 or express native nucleic acid sequences (*e.g.*, genes) that would be abnormally expressed under-expressed, or not expressed at all. The term "recombinant nucleic acid" (*e.g.*, DNA or RNA) molecule means, for example, a nucleotide sequence that is not naturally occurring or is made by the combatant (for example, artificial combination) of at least two segments of sequence that are not typically included together, not typically associated with  
25 one another, or are otherwise typically separated from one another. A recombinant nucleic acid can comprise a nucleic acid molecule formed by the joining together or combination of nucleic acid segments from different sources and/or artificially synthesized. The term "recombinantly produced" refers to an artificial combination usually accomplished by either chemical synthesis means, recursive sequence recombination of nucleic acid  
30 segments or other diversity generation methods (such as, *e.g.*, shuffling) of nucleotides, or manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques known to those of ordinary skill in the art. "Recombinantly expressed" typically refers to

techniques for the production of a recombinant nucleic acid *in vitro* and transfer of the recombinant nucleic acid into cells *in vivo*, *in vitro*, or *ex vivo* where it may be expressed or propagated. A "recombinant polypeptide" or "recombinant protein" usually refers to polypeptide or protein, respectively, that results from a cloned or recombinant gene or

5 nucleic acid.

A "subsequence" or "fragment" is any portion of an entire sequence, up to and including the complete sequence.

Numbering of a given amino acid or nucleotide polymer "corresponds to numbering" of a selected amino acid polymer or nucleic acid when the position of any

10 given polymer component (amino acid residue, incorporated nucleotide, *etc.*) is designated by reference to the same residue position in the selected amino acid or nucleotide, rather than by the actual position of the component in the given polymer.

A vector is a composition for facilitating cell transduction by a selected nucleic acid, or expression of the nucleic acid in the cell. Vectors include, *e.g.*, plasmids,

15 cosmids, viruses, YACs, bacteria, poly-lysine, *etc.* An "expression vector" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specific nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. The expression vector typically includes a nucleic acid to be transcribed operably linked to a

20 promoter.

"Substantially an entire length of a polynucleotide or amino acid sequence" refers to at least about 50%, at least about 60%, generally at least about 70%, generally at least about 80%, or typically at least about 90%, 95%, 96%, 97%, 98%, or 99% or more of a length of an amino acid sequence or nucleic acid sequence.

25 "A human alpha-interferon receptor" is a receptor which is naturally activated in human cells by an alpha interferon.

"Naturally occurring" as applied to an object refers to the fact that the object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism, including viruses, that can be isolated from a source in nature

30 and which has not been intentionally modified by man in the laboratory is naturally occurring. In one aspect, a "naturally occurring" nucleic acid (*e.g.*, DNA or RNA)

molecule is a nucleic acid molecule that exists in the same state as it exists in nature; that is, the nucleic acid molecule is not isolated, recombinant, or cloned.

As used herein, an "antibody" refers to a protein comprising one or more polypeptides substantially or partially encoded by immunoglobulin genes or fragments of immunoglobulin genes. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. A typical immunoglobulin (*e.g.*, antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (VL) and variable heavy chain (VH) refer to these light and heavy chains, respectively. Antibodies exist as intact immunoglobulins or as a number of well characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'2, a dimer of Fab which itself is a light chain joined to VH-CH1 by a disulfide bond. The F(ab)'2 may be reduced under mild conditions to break the disulfide linkage in the hinge region thereby converting the (Fab')2 dimer into an Fab' monomer. The Fab' monomer is essentially an Fab with part of the hinge region (*see Fundamental Immunology*, W.E. Paul, ed., Raven Press, N.Y. (1993), for a more detailed description of other antibody fragments). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such Fab' fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein also includes antibody fragments either produced by the modification of whole antibodies or synthesized *de novo* using recombinant DNA methodologies. Antibodies include single chain antibodies, including single chain Fv (sFv) antibodies in which a variable heavy and a variable light chain are joined together (directly or through a peptide linker) to form a continuous polypeptide.

An "antigen-binding fragment" of an antibody is a peptide or polypeptide fragment of the antibody which binds an antigen. An antigen-binding site is formed by

those amino acids of the antibody which contribute to, are involved in, or affect the binding of the antigen. *See* Scott, T.A. and Mercer, E.I., CONCISE ENCYCLOPEDIA: BIOCHEMISTRY AND MOLECULAR BIOLOGY (de Gruyter, 3d ed. 1997) [hereinafter "Scott, CONCISE ENCYCLOPEDIA"] and Watson, J.D. et al., RECOMBINANT DNA (2d ed. 1992) 5 [hereinafter "Watson, RECOMBINANT DNA"], each of which is incorporated herein by reference in its entirety for all purposes.

An "immunogen" refers to a substance that is capable of provoking an immune response. Examples of immunogens include, *e.g.*, antigens, autoantigens that play a role in induction of autoimmune diseases, and tumor-associated antigens expressed 10 on cancer cells.

An "antigen" is a substance that is capable of eliciting the formation of antibodies in a host or generating a specific population of lymphocytes reactive with that substance. Antigens are typically macromolecules (*e.g.*, proteins and polysaccharides) that are foreign to the host.

15 The term "immunoassay" includes an assay that uses an antibody or immunogen to bind or specifically bind an antigen. The immunoassay is typically characterized by the use of specific binding properties of a particular antibody to isolate, target, and /or quantify the antigen.

The term "homology" generally refers to the degree of similarity between 20 two or more structures. The term "homologous sequences" refers to regions in macromolecules that have a similar order of monomers. When used in relation to nucleic acid sequences, the term "homology" refers to the degree of similarity between two or more nucleic acid sequences (*e.g.*, genes) or fragments thereof. Typically, the degree of similarity between two or more nucleic acid sequences refers to the degree of similarity of 25 the composition, order, or arrangement of two or more nucleotide bases (or other genotypic feature) of the two or more nucleic acid sequences. The term "homologous nucleic acids" generally refers to nucleic acids comprising nucleotide sequences having a degree of similarity in nucleotide base composition, arrangement, or order. The two or more nucleic acids may be of the same or different species or group. The term "percent 30 homology" when used in relation to nucleic acid sequences, refers generally to a percent degree of similarity between the nucleotide sequences of two or more nucleic acids.

When used in relation to polypeptide (or protein) sequences, the term "homology" refers to the degree of similarity between two or more polypeptide (or protein) sequences (*e.g.*, genes) or fragments thereof. Typically, the degree of similarity between two or more polypeptide (or protein) sequences refers to the degree of similarity of the composition, order, or arrangement of two or more amino acid of the two or more polypeptides (or proteins). The two or more polypeptides (or proteins) may be of the same or different species or group. The term "percent homology" when used in relation to polypeptide (or protein) sequences, refers generally to a percent degree of similarity between the amino acid sequences of two or more polypeptide (or protein) sequences. The term "homologous polypeptides" or "homologous proteins" generally refers to polypeptides or proteins, respectively, that have amino acid sequences and functions that are similar. Such homologous polypeptides or proteins may be related by having amino acid sequences and functions that are similar, but are derived or evolved from different or the same species using the techniques described herein.

The term "subject" as used herein includes, but is not limited to, an organism; a mammal, including, *e.g.*, a human, non-human primate (*e.g.*, monkey), mouse, pig, cow, goat, rabbit, rat, guinea pig, hamster, horse, monkey, sheep, or other non-human mammal; a non-mammal, including, *e.g.*, a non-mammalian vertebrate, such as a bird (*e.g.*, a chicken or duck) or a fish; and a non-mammalian invertebrate.

The term "pharmaceutical composition" means a composition suitable for pharmaceutical use in a subject, including an animal or human. A pharmaceutical composition generally comprises an effective amount of an active agent and a pharmaceutically acceptable carrier.

The term "effective amount" means a dosage or amount sufficient to produce a desired result. The desired result may comprise an objective or subjective improvement in the recipient of the dosage or amount.

A "prophylactic treatment" is a treatment administered to a subject who does not display signs or symptoms of a disease, pathology, or medical disorder, or displays only early signs or symptoms of a disease, pathology, or disorder, such that treatment is administered for the purpose of diminishing, preventing, or decreasing the risk of developing the disease, pathology, or medical disorder. A prophylactic treatment functions as a preventative treatment against a disease or disorder. A "prophylactic

"activity" is an activity of an agent, such as a nucleic acid, vector, gene, polypeptide, protein, substance, composition thereof that, when administered to a subject who does not display signs or symptoms of pathology, disease or disorder, or who displays only early signs or symptoms of pathology, disease, or disorder, diminishes, prevents, or decreases 5 the risk of the subject developing a pathology, disease, or disorder. A "prophylactically useful" agent or compound (*e.g.*, nucleic acid or polypeptide) refers to an agent or compound that is useful in diminishing, preventing, treating, or decreasing development of pathology, disease or disorder.

A "therapeutic treatment" is a treatment administered to a subject who 10 displays symptoms or signs of pathology, disease, or disorder, in which treatment is administered to the subject for the purpose of diminishing or eliminating those signs or symptoms of pathology, disease, or disorder. A "therapeutic activity" is an activity of an agent, such as a nucleic acid, vector, gene, polypeptide, protein, substance, or composition thereof, that eliminates or diminishes signs or symptoms of pathology, disease or disorder, 15 diminishes when administered to a subject suffering from such signs or symptoms. A "therapeutically useful" agent or compound (*e.g.*, nucleic acid or polypeptide) indicates that an agent or compound is useful in diminishing, treating, or eliminating such signs or symptoms of a pathology, disease or disorder.

The term "gene" broadly refers to any segment of DNA associated with a 20 biological function. Genes include coding sequences and/or regulatory sequences required for their expression. Genes also include non-expressed DNA nucleic acid segments that, *e.g.*, form recognition sequences for other proteins.

Generally, the nomenclature used hereafter and the laboratory procedures 25 in cell culture, molecular genetics, molecular biology, nucleic acid chemistry, and protein chemistry described below are those well known and commonly employed by those of ordinary skill in the art. Standard techniques, such as described in Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 (hereinafter "Sambrook") and *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, a joint 30 venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (supplemented through 1999) (hereinafter "Ausubel"), are used for recombinant nucleic acid methods, nucleic acid synthesis, cell culture methods, and transgene incorporation,

e.g., electroporation, injection, and lipofection. Generally, oligonucleotide synthesis and purification steps are performed according to specifications. The techniques and procedures are generally performed according to conventional methods in the art and various general references which are provided throughout this document. The procedures 5 therein are believed to be well known to those of ordinary skill in the art and are provided for the convenience of the reader.

A variety of additional terms are defined or otherwise characterized herein.

#### POLYNUCLEOTIDES OF THE INVENTION

##### Interferon-alpha Homologue Sequences

10 The invention provides isolated or recombinant interferon-alpha homologue polypeptides, and isolated or recombinant polynucleotides encoding the polypeptides.

As described in more detail below, in accordance with the present invention, polynucleotide sequences which encode novel interferon-alpha homologue polypeptides, nucleotide sequences (e.g., subsequences) that encode fragments of 15 interferon-alpha homologue polypeptides, and nucleotide sequences that encode related fusion polypeptides or proteins, or functional equivalents thereof, are collectively referred to herein as "interferon-alpha homologues," "interferon homologue nucleic acids," "IFN-alpha homologues," "IFN homologues," "IFN nucleic acids," "interferon homologues," "interferon nucleic acids," "recombinant interferon-alpha," "recombinant interferon-alpha 20 nucleic acids," "nucleic acids of the invention," "polynucleotides of the invention," or "nucleotides of the invention." Polynucleotide, nucleotide are nucleic acid fragments of each of the preceding terms are also intended to be included and encompassed in polynucleotides, nucleotides, and nucleic acids of the invention. The term "nucleic acid" is used interchangeable with the term "nucleotide."

25 Polynucleotides encoding the polypeptides of the invention were discovered in libraries of shuffled interferon-alpha related sequences. The library members were screened for antiproliferative activity against human tumor cell lines and, in some cases, assayed for antiviral activity against virus-infected human cells. A subset of the sequences provided herein were discovered in shuffled libraries screened for 30 antiviral activity against virus-infected mouse cells. Coding sequences for interferon homologues were identified as described in the examples.

Briefly, libraries of shuffled mature interferon-alpha coding sequences were introduced into *E. coli*. Colonies were screened in a high-throughput antiproliferative activity assay against a human Daudi tumor cell line as described in Example 1, and colonies expressing active polypeptides were selected, re-screened, and expression levels 5 determined. DNA from selected colonies was isolated and re-shuffled to create secondary libraries. The secondary libraries were introduced into *E. coli* and screened for antiproliferative activity in the human Daudi cell line-based cell proliferation assay. DNA from colonies selected from the primary and secondary library screens were transduced into Chinese hamster ovary (CHO) cells, and stable cell lines were generated. CHO- 10 expressed proteins were purified, quantitated, and assayed for antiproliferative activity using the human Daudi cell line, and optionally, for antiviral activity using encephalomyocarditis virus (EMCV)-infected human WISH cells, as described in Example 1. Exemplary shuffled nucleic acids which encode interferon-alpha homologue polypeptides having antiproliferative activity in the human Daudi cell line-based assay are 15 identified herein as SEQ ID NO:1 to SEQ ID NO:35, which encode mature interferon-alpha homologue polypeptides identified herein as SEQ ID NO:36 to SEQ ID NO:70, respectively. Libraries of shuffled mature interferon-alpha coding sequences were also screened in a high-throughput antiviral activity screen against EMCV-infected mouse cells. Exemplary shuffled nucleic acids which encode polypeptides having antiviral 20 activity in the murine cell /EMCV-based assay are identified herein as SEQ ID NO:72 to SEQ ID NO:78, which encode mature interferon homologue polypeptides identified herein as SEQ ID NO:79 to SEQ ID NO:85.

In another aspect, the invention provides an isolated or recombinant nucleic acid that comprises a polynucleotide sequence selected from the group of: (a) SEQ ID 25 NO:1 to SEQ ID NO:35, or a complementary polynucleotide sequence thereof; (b) a polynucleotide sequence encoding a polypeptide selected from SEQ ID NO:36 to SEQ ID NO:71, or a complementary polynucleotide sequence thereof; (c) a polynucleotide sequence which hybridizes under at least stringent or at least highly stringent hybridization conditions (or ultra-high stringent or ultra-ultra- high stringent hybridization conditions) 30 over substantially the entire length of polynucleotide sequence (a) or (b), or with a 50, 120, 130, 140, 145, 150, 155, 160, or 165 nucleotide base subsequence or fragment of a polynucleotide sequence of (a) or (b); and (d) a polynucleotide sequence comprising a

fragment of (a), (b), or (c), which fragment encodes all or a part of a polypeptide having an antiproliferative activity in a human Daudi cell line-based assay or an antiviral activity in an assay known in the art for measuring antiviral activity.

In another aspect, the invention provides an isolated or recombinant nucleic acid that comprises a polynucleotide sequence selected from the group of: (a) SEQ ID NO:72 to SEQ ID NO:78, or a complementary polynucleotide sequence thereof; (b) a polynucleotide sequence encoding a polypeptide selected from SEQ ID NO:79 to SEQ ID NO:85, or a complementary polynucleotide sequence thereof; (c) a polynucleotide sequence which hybridizes under at least stringent or at least highly stringent hybridization conditions (or ultra-high stringent or ultra-ultra- high stringent hybridization conditions) over substantially the entire length of polynucleotide sequence (a) or (b), or with a 50, 120, 130, 140, 145, 150, 155, 160, or 165 nucleotide base subsequence or fragment of a polynucleotide sequence of (a) or (b); and (d) a polynucleotide sequence comprising a fragment of (a), (b), or (c), which fragment encodes all or a part of a polypeptide having an antiproliferative activity in a human Daudi cell line-based assay or an antiviral activity in a murine cell line/EMCV-based assay.

The present invention also includes a mature interferon-alpha homologue polypeptide comprising the amino acid identified herein as SEQ ID NO:71 and a polynucleotide sequence encoding said polypeptide or a fragment of said polypeptide having an antiproliferative activity in the human Daudi cell line-based assay and/or an antiviral activity in the murine cell /EMCV-based assay.

The invention also includes an isolated or recombinant nucleic acid comprising a polynucleotide sequence encoding a polypeptide, wherein the polypeptide comprises the amino acid sequence: CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QQLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof, where X<sub>11</sub> is N or D; X<sub>12</sub> is R, S, or K; X<sub>15</sub> is L or M; X<sub>16</sub> is I, M, or V; X<sub>19</sub> is A or G; X<sub>22</sub> is G or R; X<sub>24</sub> is I or T; X<sub>26</sub> is P or H; X<sub>34</sub> is H, Y or Q; X<sub>38</sub> is F or L; X<sub>40</sub> is Q or R; X<sub>45</sub> is G or S; X<sub>46</sub> is N or H; X<sub>47</sub> is Q or R; X<sub>50</sub> is K or R; X<sub>51</sub> is A or T; X<sub>55</sub> is S or F; X<sub>56</sub>

is V or A; X<sub>57</sub> is L or F; X<sub>60</sub> is M or I; X<sub>61</sub> is I or M; X<sub>67</sub> is L or F; X<sub>72</sub> is D or N; X<sub>75</sub> is A or V; X<sub>76</sub> is A or T; X<sub>78</sub> is E or D; X<sub>79</sub> is Q or E; X<sub>80</sub> is S, R, T, or N; X<sub>83</sub> is E or D; X<sub>85</sub> is F or L; X<sub>86</sub> is S or Y; X<sub>88</sub> is E or G; X<sub>90</sub> is Y, H, N; X<sub>95</sub> is D, E, or N; X<sub>101</sub> is I, M, or V; X<sub>103</sub> is E or G; X<sub>105</sub> is G or W; X<sub>106</sub> is V or M; X<sub>107</sub> is E, G, or K; X<sub>108</sub> is E or G; X<sub>114</sub> is V, 5 E, or G; X<sub>116</sub> is S or P; X<sub>121</sub> is K or R; X<sub>124</sub> is F or L; X<sub>132</sub> is T, I, or M; X<sub>134</sub> is K or R; and X<sub>140</sub> is A or S. Each of the single letters of this amino acid sequence represents a 10 particular amino acid residue according to standard practice known to those of ordinary skill in the art. Such polypeptides having an antiproliferative activity in the human Daudi cell line-based assay (*e.g.*, at least about  $8.3 \times 10^6$  units/mg) and/or an antiviral activities in a human WISH cell/EMCV-based assay (at least about  $2.1 \times 10^7$  units/mg).

As described in greater detail below, the polynucleotides of the invention are useful in for a variety of applications, including, but not limited to, as therapeutic and prophylactic agents in methods of *in vivo* and *ex vivo* treatment of a variety of diseases, disorders, and conditions in a variety of subjects; for use in *in vitro* methods, such as 15 diagnostic methods, to detect, diagnose, and treat a variety of diseases, disorders, and conditions in a variety of subjects; for use in, *e.g.*, gene therapy; as therapeutics and prophylactics, *e.g.*, for use in methods of therapeutic and prophylactic treatment of a disease, disorder or condition; as immunogens; for use in diagnostic and screening assays; and as diagnostic probes for the presence of complementary or partially complementary 20 nucleic acids (including for detection of IFN-alpha coding nucleic acids).

#### Making Polynucleotides of the Invention

Polynucleotides and oligonucleotides of the invention can be prepared by standard solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 20, 30, 40, 50, 60, 70, 80, 90, and/or 100 nucleotide bases are 25 individually synthesized, then joined (*e.g.*, by enzymatic or chemical ligation methods, or polymerase mediated recombination methods) to form essentially any desired continuous sequence. In another aspect, nucleotide fragments of greater than 100 nucleotide bases (*e.g.*, 150, 180, 200, 210, 240, 270, 300, 330, 360, 390, 400, 420, 450, 465, 474, 470, 475, 489, 490, 495, 496 bases) are individually synthesized, then joined (*e.g.*, by enzymatic or 30 chemical ligation methods, or polymerase mediated recombination methods) to form essentially any desired continuous sequence. example, the polynucleotides and oligonucleotides of the invention, including fragments thereof (and those as described

herein), can be prepared by chemical synthesis using, e.g., the classical phosphoramidite method described by Beaucage *et al.* (1981) *Tetrahedron Letters* 22:1859-69, or the method described by Matthes *et al.* (1984) *EMBO J.* 3:801-05., e.g., as is typically practiced in automated synthetic methods. According to the phosphoramidite method, 5 oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

In addition, essentially any nucleic acid can be custom ordered from any of a variety of commercial sources, such as The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (<http://www.genco.com>), 10 ExpressGen Inc. ([www.expressgen.com](http://www.expressgen.com)), Operon Technologies Inc. (Alameda, CA) and many others. Similarly, peptides and antibodies can be custom ordered from any of a variety of sources, such as PeptidoGenic (pkim@ccnet.com), HTI Bio-products, inc. (<http://www.htibio.com>), BMA Biomedicals Ltd. (U.K.), Bio.Synthesis, Inc., and many others.

15 Certain polynucleotides of the invention may also obtained by screening cDNA libraries (e.g., libraries generated by recombining homologous nucleic acids as in typical diversity generation methods, such as, e.g., shuffling methods) using oligonucleotide probes which can hybridize to or PCR-amplify polynucleotides which encode the interferon homologue polypeptides and fragments of those polypeptides. 20 Procedures for screening and isolating cDNA clones are well-known to those of skill in the art. Such techniques are described in, for example, Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 (hereinafter "Sambrook") and *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between 25 Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (supplemented through 1999) (hereinafter "Ausubel").

As described in more detail herein, the polynucleotides of the invention include sequences which encode novel mature interferon-alpha homologues and sequences complementary to the coding sequences, and novel fragments of such coding sequences 30 and complements thereof. The polynucleotides can be in the form of RNA or in the form of DNA, and include mRNA, cRNA, synthetic RNA and DNA, and cDNA. The polynucleotides can be double-stranded or single-stranded, and if single-stranded, can be

the coding strand or the non-coding (anti-sense, complementary) strand. The polynucleotides optionally include the coding sequence of an interferon-alpha homologue (i) in isolation, (ii) in combination with additional coding sequence, so as to encode, e.g., a fusion protein, a pre-protein, a prepro-protein, or the like, (iii) in combination with non-coding sequences, such as introns, control elements such as a promoter, a terminator element, or 5' and/or 3' untranslated regions effective for expression of the coding sequence in a suitable host, and/or (iv) in a vector or host environment in which the interferon-alpha homologue coding sequence is a heterologous nucleic acid sequence or gene. Sequences can also be found in combination with typical compositional formulations of nucleic acids, including in the presence of carriers, buffers, adjuvants, excipients and the like.

The term DNA or RNA encoding the respective interferon-alpha homologue polypeptide includes any oligodeoxynucleotide or oligodeoxyribonucleotide sequence which, upon expression in an appropriate host cell, results in production of an interferon-alpha homologue polypeptide of the invention. The DNA or RNA can be produced in an appropriate host cell, or in a cell-free (*in vitro*) system, or can be produced synthetically (e.g., by an amplification technique such as PCR) or chemically.

#### Using Polynucleotides of the Invention

The polynucleotides of the invention have a variety of uses in, for example: recombinant production (*i.e.*, expression) of the interferon-alpha homologue polypeptides of the invention; as therapeutics and prophylactics, *e.g.*, for use in methods of therapeutic and prophylactic treatment of a disease, disorder or condition; for use in, gene therapy methods and related applications;; as immunogens; for use in diagnostic and screening assays; as diagnostic probes for the presence of complementary or partially complementary nucleic acids (including for detection of natural IFN- alpha coding nucleic acids); as substrates for further reactions, *e.g.*, shuffling reactions or mutation reactions to produce new and/or improved IFN-alpha homologues, and the like.

#### EXPRESSION OF POLYPEPTIDES

In accordance with the present invention, polynucleotide sequences which encode novel and/or mature interferon-alpha homologues, fragments of interferon-alpha proteins, related fusion proteins, or functional equivalents thereof, are collectively referred to herein as "interferon-alpha homologue polypeptides," "interferon-alpha homologue

proteins," or "interferon-alpha homologues," "interferon homologues," "IFN-alpha homologues," "IFN homologues", "IFN polypeptides," "IFN proteins" "polypeptides of the invention," or "proteins of the invention." Polypeptide or amino acid fragments of each of the preceding terms are also intended to be included and encompassed in the polypeptides or proteins of the invention. Such polynucleotide sequences of the invention are used in recombinant DNA (or RNA) molecules that direct the expression of the interferon-alpha homologue polypeptides in appropriate host cells. Due to the inherent degeneracy of the genetic code, other nucleic acid sequences which encode substantially the same or a functionally equivalent amino acid sequence are also used to clone and express the interferon homologues.

Modified Coding Sequences

As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence (including, e.g., a nucleotide sequence encoding an interferon-alpha homologue of the invention or a fragment thereof) to enhance its expression in a particular host. The genetic code is redundant with 64 possible codons, but most organisms preferentially use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons (see, e.g., Zhang S.P. et al. (1991) *Gene* 105:61-72). Codons can be substituted to reflect the preferred codon usage of the host, a process called "codon optimization" or "controlling for species codon bias."

Optimized coding sequence containing codons preferred by a particular prokaryotic or eukaryotic host (see also Murray, E. et al. (1989) *Nuc. Acids Res.* 17:477-508) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, preferred stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The preferred stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* prefer to use UAA as the stop codon (Dalphin M.E. et al. (1996) *Nuc. Acids Res.* 24:216-218).

The polynucleotide sequences of the present invention can be engineered in order to alter an interferon homologue coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of

the gene product. For example, alterations may be introduced using techniques which are well known in the art, *e.g.*, site-directed mutagenesis, to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to introduce splice sites, *etc.*

5 Vectors, Promoters and Expression Systems

The present invention also includes recombinant constructs comprising one or more of the nucleic acid sequences as broadly described herein (*e.g.*, those encoding an interferon-alpha homologue of the invention or a fragment thereof). The constructs comprise a vector, such as, a plasmid, a cosmid, a phage, a virus (including a retrovirus), a bacterial artificial chromosome (BAC), a yeast artificial chromosome (YAC), and the like, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available.

10 General texts which describe molecular biological techniques useful herein, including the use of vectors, promoters and many other relevant topics, include Juo, P-S., CONCISE DICTIONARY OF BIOMEDICAL AND MOLECULAR BIOLOGY (CRC Press 1996); Singleton *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY (2d ed. 1994); THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY (Walker ed., 1988); Hale & 15 Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY (1991); Scott and Mercer, CONCISE ENCYCLOPEDIA OF BIOCHEMISTRY AND MOLECULAR BIOLOGY (3d ed. 1997); Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, volume 152 Academic Press, Inc., San Diego, CA (hereinafter "Berger"); Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor 20 Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") and *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (supplemented through 1999) ("Ausubel")). Examples of techniques sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR), the ligase 25 chain reaction (LCR), Q $\beta$ -replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA), *e.g.*, for the production of the homologous nucleic acids of the invention are found in Berger, Sambrook, and Ausubel, as well as Mullis *et al.* (1987) U.S. 30

Patent No. 4,683,202; U.S. Pat. No. 4,683,195, issued July 28, 1997; *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, eds.) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (October 1, 1990) *C&EN* 36-47; *The Journal Of NIH Research* (1991) 3, 81-94; (Kwoh *et al.* (1989) *Proc. Nat'l Acad. Sci. USA* 86, 1173; 5 Guatelli *et al.* (1990) *Proc. Nat'l Acad. Sci. USA* 87, 1874; Lomell *et al.* (1989) *J. Clin. Chem.* 35, 1826; Landegren *et al.* (1988) *Science* 241, 1077-1080; Van Brunt (1990) *Biotechnology* 8, 291-294; Wu and Wallace (1989) *Gene* 4, 560; Barringer *et al.* (1990) *Gene* 89, 117, and Sooknanan and Malek (1995) *Biotechnology* 13:563-564.

PCR generally refers to a procedure wherein minute amounts of a specific 10 piece of nucleic acid, RNA, and/or DNA, are amplified by methods well known in the art (see, e.g., U.S. Pat. No. 4,683,195 and other references above). Generally, sequence information from the ends of the region of interest or beyond is used, for design of oligonucleotide primers. Such primers will be identical or similar in sequence to the opposite strands of the template to be amplified. The 5' terminal nucleotides of the 15 opposite strands may coincide with the ends of the amplified material. PCR may be used to amplify specific RNA or specific DNA sequences, recombinant DNA or RNA sequences, DNA and RNA sequences from total genomic DNA, and cDNA transcribed from total cellular RNA, bacteriophage or plasmid sequences, *etc.* PCR is one example, but not the only example, of a nucleic acid polymerase reaction method for amplifying a 20 nucleic acid test sample comprising the use of another (e.g., known) nucleic acid as a primer. Improved methods of cloning *in vitro* amplified nucleic acids are described in Wallace *et al.*, U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng *et al.* (1994) *Nature* 369:684-685 and the 25 references therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See Ausubel, Sambrook and Berger, all supra.*

The present invention also relates to host cells which are transduced with 30 vectors of the invention, and the production of polypeptides of the invention (including fragments thereof) by recombinant techniques. Host cells are genetically engineered (*i.e.*, transduced, transformed or transfected) with the vectors of this invention, which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in

the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the interferon homologue gene. The culture conditions, such as temperature, pH and the like, are those previously used with the host 5 cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, e.g., Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, 3d ed., Wiley-Liss, New York and the references cited therein.

The interferon homologue polypeptides and proteins of the invention can 10 also be produced in non-animal cells such as plants, yeast, fungi, bacteria and the like. In addition to Sambrook, Berger and Ausubel, details regarding cell culture can be found in Payne et al. (1992) *Plant Cell and Tissue Culture in Liquid Systems*, John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds.) (1995) *Plant Cell, Tissue and Organ Culture; Fundamental Methods*, Springer Lab Manual, Springer-Verlag (Berlin 15 Heidelberg New York) and Atlas and Parks (eds.) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

The polynucleotides of the present invention may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; 20 bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, pseudorabies, adenovirus, adeno-associated virus, retroviruses and many others. Any vector that transduces genetic material into a cell, and, if replication is desired, which is replicable and viable in the relevant host can be used.

25 The nucleic acid sequence in the expression vector is operatively linked to an appropriate transcription control sequence (promoter) to direct mRNA synthesis. Examples of such promoters include: LTR or SV40 promoter, *E. coli* lac or trp promoter, phage lambda P<sub>L</sub> promoter, and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a 30 ribosome binding site for translation initiation, and a transcription terminator. The vector optionally includes appropriate sequences for amplifying expression. In addition, the expression vectors optionally comprise one or more selectable marker genes to provide a

phenotypic trait for selection of transformed host cells, such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

The vector containing the appropriate DNA sequence as described herein, 5 as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. Examples of appropriate expression hosts include: bacterial cells, such as *E. coli*, *Streptomyces*, and *Salmonella typhimurium*; fungal cells, such as *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Neurospora crassa*; insect cells such as *Drosophila* and *Spodoptera frugiperda*; 10 mammalian cells such as CHO, COS, BHK, HEK 293 or Bowes melanoma; plant cells, etc. It is understood that not all cells or cell lines need to be capable of producing fully functional interferon homologues; for example, antigenic fragments of an interferon homologue may be produced in a bacterial or other expression system. The invention is not limited by the host cells employed.

15 In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the interferon homologue. For example, when large quantities of interferon homologue or fragments thereof are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the 20 interferon homologue coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke & Schuster (1989) *J. Biol. Chem.* 264:5503-5509); pET vectors (Novagen, Madison WI); and the like.

25 Similarly, in the yeast *Saccharomyces cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used for production of the interferon homologue proteins of the invention. For reviews, see Ausubel *et al.* (*supra*) and Grant *et al.* (1987; *Methods in Enzymology* 153:516-544).

30 In mammalian host cells, a number expression systems, such as viral-based systems, may be utilized. In cases where an adenovirus is used as an expression vector, a coding sequence is optionally ligated into an adenovirus transcription/translation complex

consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing interferon homologue in infected host cells (Logan and Shenk (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the rous sarcoma virus 5 (RSV) enhancer, may be used to increase expression in mammalian host cells.

**Additional Expression Elements**

Specific initiation signals can aid in efficient translation of an interferon homologue coding sequence. These signals can include, e.g., the ATG initiation codon and adjacent sequences. In cases where interferon homologue coding sequence, its 10 initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence (e.g., a mature protein coding sequence), or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame 15 to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-62; Bittner *et al.* (1987) *Methods in Enzymol.* 153:516-544).

20 **Secretion/Localization Sequences**

Polynucleotides of the invention can also be fused, for example, in-frame to nucleic acid encoding a secretion/localization sequence, to target polypeptide expression to a desired cellular compartment, membrane, or organelle, or to direct polypeptide secretion to the periplasmic space or into the cell culture media. Such sequences are 25 known to those of skill, and include secretion leader peptides, organelle targeting sequences (e.g., nuclear localization sequences, ER retention signals, mitochondrial transit sequences, chloroplast transit sequences), membrane localization/anchor sequences (e.g., stop transfer sequences, GPI anchor sequences), and the like. Polypeptides expressed by such polynucleotides of the invention may include the amino acid sequence corresponding 30 to the secretion and/or localization sequence(s).

Expression Hosts

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a eukaryotic cell, such as a mammalian cell, a yeast cell, or a plant cell, or the host cell can be a prokaryotic cell,

5 such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, electroporation, or other common techniques (Davis, L., Dibner, M., and Battey, J. (1986) *Basic Methods in Molecular Biology*). The cell may include a nucleic acid of the invention, said nucleic acid encoding a polypeptide, wherein said cells expresses a polypeptide (e.g., an

10 interferon-alpha homologue polypeptide having an antiviral or anti-proliferative activity as measured by the assays described herein). The invention also includes a vector comprising any nucleic acid of the invention described herein and includes a cell transduced by such a vector. Furthermore, Cells and transgenic animals which include any polypeptide or nucleic acid above or throughout this specification, e.g., produced by

15 transduction of a vector of the invention, are an additional feature of the invention.

A host cell strain is optionally chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a "pre" or a "pro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, BHK, MDCK, 293, WI38, etc. have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

25 For long-term, high-yield production of recombinant proteins, stable expression can be used. For example, cell lines which stably express a polypeptide of the invention are transduced using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. For example, resistant clumps of

stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Host cells transformed with a nucleotide sequence encoding a polypeptide of the invention are optionally cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein or fragment thereof produced by a recombinant cell may be secreted, membrane-bound, or contained intracellularly, depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides encoding mature interferon homologues of the invention can be designed with signal sequences which direct secretion of the mature polypeptides through a prokaryotic or eukaryotic cell membrane.

Additional Polypeptide Sequences

The polynucleotides of the present invention may also comprise a coding sequence fused in-frame to a marker sequence which, *e.g.*, facilitates purification of the encoded polypeptide of the invention. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, a sequence which binds glutathione (*e.g.*, GST), a hemagglutinin (HA) tag (corresponding to an epitope derived from the influenza hemagglutinin protein; Wilson, I. *et al.* (1984) *Cell* 37:767), maltose binding protein sequences, the FLAG epitope utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA), and the like. The inclusion of a protease-cleavable polypeptide linker sequence between the purification domain and the interferon homologue sequence is useful to facilitate purification. One expression vector contemplated for use in the compositions and methods described herein provides for expression of a fusion protein comprising a polypeptide of the invention fused to a polyhistidine region separated by an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography, as described in Porath *et al.* (1992) *Protein Expression and Purification* 3:263-281), while the enterokinase cleavage site provides a means for separating the interferon homologue polypeptide from the fusion protein. pGEX vectors (Promega; Madison, WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed

cells by adsorption to ligand-agarose beads (e.g., glutathione-agarose in the case of GST-fusions) followed by elution in the presence of free ligand.

Polypeptide Production and Recovery

Following transduction of a suitable host strain and growth of the host

- 5 strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including  
10 freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other methods, which are well known to those skilled in the art.

As noted, many references are available for the culture and production of many cells, including cells of bacterial, plant, animal (especially mammalian) and archebacterial origin. See, e.g., Sambrook, Ausubel, and Berger (all *supra*), as well as  
15 Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, third edition, Wiley- Liss, New York and the references cited therein; Doyle and Griffiths (1997) *Mammalian Cell Culture: Essential Techniques*, John Wiley and Sons, NY; Humason (1979) *Animal Tissue Techniques*, 4th edition, W.H. Freeman and Company; and Ricciardelli *et al.* (1989) *In vitro Cell Dev. Biol.* 25:1016-1024. For plant cell culture and  
20 regeneration, Payne *et al.* (1992) *Plant Cell and Tissue Culture in Liquid Systems*, John Wiley & Sons, Inc., New York, NY; Gamborg and Phillips (eds.) (1995) *Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual*, Springer-Verlag (Berlin Heidelberg New York) and *Plant Molecular Biology* (1993) R.R.D. Croy, ed., Bios Scientific Publishers, Oxford, U.K. ISBN 0 12 198370 6. Cell culture media in  
25 general are set forth in Atlas and Parks (eds.) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL. Additional information for cell culture is found in available commercial literature such as the *Life Science Research Cell Culture Catalogue* (1998) from Sigma- Aldrich, Inc (St. Louis, MO) ("Sigma-LSRCCC") and, e.g., the *Plant Culture Catalogue* and supplement (1997) also from Sigma-Aldrich, Inc. (St. Louis, MO) ("Sigma-PCCS").

Polypeptides of the invention can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including

ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography (*e.g.*, using any of the tagging systems noted herein), hydroxyapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as desired, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed in the final purification steps. In addition to the references noted *supra*, a variety of purification methods are well known in the art, including, *e.g.*, those set forth in Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; and Bollag et al. (1996) *Protein Methods*, 2<sup>nd</sup> Edition, Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook*, Humana Press, NJ, Harris and Angal (1990) *Protein Purification Applications: A Practical Approach*, IRL Press at Oxford, Oxford, England; Harris and Angal, *Protein Purification Methods: A Practical Approach*, IRL Press at Oxford, Oxford, England; Scopes (1993) *Protein Purification: Principles and Practice* 3<sup>rd</sup> Edition, Springer Verlag, NY; Janson and Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications*, Second Edition, Wiley-VCH, NY; and Walker (1998) *Protein Protocols on CD-ROM*, Humana Press, NJ.

In vitro Expression Systems

Cell-free transcription/translation systems can also be employed to produce polypeptides using DNAs or RNAs of the present invention. Several such systems are commercially available. A general guide to *in vitro* transcription and translation protocols is found in Tymms (1995) *In vitro Transcription and Translation Protocols: Methods in Molecular Biology*, Volume 37, Garland Publishing, NY.

Modified Amino Acids

Polypeptides of the invention may contain one or more modified amino acids. The presence of modified amino acids may be advantageous in, for example, (a) increasing polypeptide serum half-life, (b) reducing polypeptide antigenicity, (c) increasing polypeptide storage stability. Amino acid(s) are modified, for example, co-translationally or post-translationally during recombinant production (*e.g.*, N-linked glycosylation at N-X-S/T motifs during expression in mammalian cells) or modified by synthetic means.

Non-limiting examples of a modified amino acid include a glycosylated amino acid, a sulfated amino acid, a prenylated (*e.g.*, farnesylated, geranylgeranylated) amino acid, an acetylated amino acid, an acylated amino acid, a PEG-ylated amino acid, a biotinylated amino acid, a carboxylated amino acid, a phosphorylated amino acid, and the like. References adequate to guide one of skill in the modification of amino acids are replete throughout the literature. Example protocols are found in Walker (1998) *Protein Protocols* on CD-ROM Human Press, Towata, NJ.

The polynucleotides and polypeptides of the invention have a variety of uses, including, but not limited to, for example: in recombinant production (*i.e.*, expression) of the recombinant interferon-alpha homologues of the invention; as therapeutic and prophylactic agents in methods of *in vivo* and *ex vivo* treatment of a variety of diseases, disorders, and conditions in a variety of subjects; for use in *in vitro* methods, such as diagnostic and screening methods, to detect, diagnose, and treat a variety of diseases, disorders, and conditions (*e.g.*, cancers, viral-based disorders, angiogenic-based disorders) in a variety of subjects (*e.g.*, mammals); as immunogens; in gene therapy methods and DNA- or RNA-based delivery methods to deliver or administer *in vivo*, *ex vivo*, or *in vitro* biologically active polypeptides of the invention to a tissue, population or cells, organ, graft, bodily system of a subject (*e.g.*, organ system, lymphatic system, blood system, *etc.*); as DNA vaccines, multi-component vaccines for use in prophylactic or therapeutic treatment of a variety of diseases, disorders, or other conditions (*e.g.*, cancers, viral-based disorders, angiogenic-based disorders) in a variety of subjects (*e.g.*, mammals); as adjuvants to enhance or augment an immune response in a subject; as a component of a multiple-step boosting vaccination method (*e.g.*, a format comprising a prime vaccination by delivery of a DNA or RNA nucleotide (*e.g.*, a nucleotide encoding a polypeptide of the invention or encoding another polypeptide) followed by a second boost of a polypeptide (*e.g.*, a polypeptide of the invention or other polypeptide); as diagnostic probes for the presence of complementary or partially complementary nucleic acids (including for detection of natural interferon-alpha coding nucleic acids); as substrates for further reactions, *e.g.*, shuffling reactions, mutation reactions, or other diversity generation reactions to produce new and/or improved interferon-alpha homologues and new interferon-alpha nucleic acids encoding such homologues, *e.g.*, to evolve novel therapeutic or prophylactic properties, and the like; for polymerase chain reactions (PCR) or cloning

methods, *e.g.*, including digestion or ligation reactions, to identify new and/or improved naturally-occurring or non-naturally occurring IFN-alpha nucleic acids and polypeptides encoded therefrom. Polynucleotides which encode an interferon homologue of the invention, or complements of the polynucleotides, are optionally administered to a cell to

5 accomplish a therapeutically or prophylactically useful process or to express a therapeutically useful product *in vivo*, *ex vivo*, or *in vitro*. These applications, including *in vivo* or *ex vivo* applications, including, *e.g.*, gene therapy, include a multitude of techniques by which gene expression may be altered in cells. Such methods include, for instance, the introduction of genes for expression of, *e.g.*, therapeutically or

10 prophylactically useful polypeptides, such as the interferon homologues of the present invention. Such methods include, for example, infecting with a retrovirus comprising the polynucleotides and/or polypeptides of the invention. Optionally, the retrovirus further comprises additional exogenous, *e.g.*, therapeutic or prophylactic gene construct, sequences. In one aspect, the invention provides gene therapy methods of prophylactically

15 or therapeutically treating a disease, disorder or condition in a subject in need of such treatment by administering *in vivo*, *ex vivo*, or *in vitro* one or more nucleic acids of the invention described herein to one or more cells of a subject, including an organism or mammal, including, *e.g.*, a human, primate, mouse, pig, cow, goat, rabbit, rat, guinea pig, hamster, horse, sheep; or a non-mammalian vertebrate such as a bird (*e.g.*, a chicken or

20 duck) or a fish, or invertebrate, as described in more detail below.

In another aspect, the invention provides methods of prophylactically or therapeutically treating a disease, disorder or condition in a subject in need of such treatment by administering *in vivo*, *ex vivo*, or *in vitro* one or more polypeptides of the invention described herein to one or more cells of a subject (including those defined herein), as described in more detail below.

#### Polypeptide Expression

Polynucleotides encoding interferon homologue polypeptides of the invention are particularly useful for *in vivo* or *ex vivo* therapeutic or prophylactic applications, using techniques well known to those skilled in the art. For example,

30 cultured cells are engineered *ex vivo* with a polynucleotide (DNA or RNA), with the engineered cells then being returned to the patient. Cells may also be engineered *in vivo* or *ex vivo* for expression of a polypeptide *in vivo* or *ex vivo*, respectively.

A number of viral vectors suitable for organismal *in vivo* or *ex vivo* transduction and expression are known. Such vectors include retroviral vectors (see Miller(1992) *Curr. Top. Microbiol. Immunol.* 158:1-24; Salmons and Gunzburg (1993) *Human Gene Therapy* 4:129-141; Miller *et al.* (1994) *Methods in Enzymology* 217:581-599) and adeno-associated vectors (reviewed in Carter (1992) *Curr. Opinion Biotech.* 3:533-539; Muzyczka (1992) *Curr. Top. Microbiol. Immunol.* 158:97-129). Other viral vectors that are used include adenoviral vectors, herpes viral vectors and Sindbis viral vectors, as generally described in, e.g., Jolly (1994) *Cancer Gene Therapy* 1:51-64; Latchman (1994) *Molec. Biotechnol.* 2:179-195; and Johanning *et al.* (1995) *Nucl. Acids Res.* 23:1495-1501.

Gene therapy provides methods for combating chronic infectious diseases (e.g., HIV infection, viral hepatitis, Herpes Simplex Virus (HSV), hepatitis B (HepB), dengue virus, etc.), as well as non-infectious diseases including cancer and allergic diseases and some forms of congenital defects such as enzyme deficiencies. Several approaches for introducing nucleic acids into cells *in vivo*, *ex vivo* and *in vitro* have been used. These include liposome based gene delivery (Debs and Zhu (1993) WO 93/24640 and U.S. Pat. No. 5,641,662; Mannino and Gould-Fogerite (1988) *BioTechniques* 6(7):682-691; Rose, U.S. Pat No. 5,279,833; Brigham (1991) WO 91/06309; and Felgner *et al.* (1987) *Proc. Nat'l Acad. Sci. USA* 84:7413-7414); Brigham *et al.* (1989) *Am. J. Med. Sci.* 298:278-281; Nabel *et al.* (1990) *Science* 249:1285-1288; Hazinski *et al.* (1991) *Am. J. Resp. Cell Molec. Biol.* 4:206-209; and Wang and Huang (1987) *Proc. Nat'l Acad. Sci. (USA)* 84:7851-7855); adenoviral vector mediated gene delivery, e.g., to treat cancer (see, e.g., Chen *et al.* (1994) *Proc. Nat'l Acad. Sci. USA* 91:3054-3057; Tong *et al.* (1996) *Gynecol. Oncol.* 61:175-179; Clayman *et al.* (1995) *Cancer Res.* 5:1-6; O'Malley *et al.* (1995) *Cancer Res.* 55:1080-1085; Hwang *et al.* (1995) *Am. J. Respir. Cell Mol. Biol.* 13:7-16; Haddada *et al.* (1995) *Curr. Top. Microbiol. Immunol.* 199 (Pt. 3):297-306; Addison *et al.* (1995) *Proc. Nat'l Acad. Sci. USA* 92:8522-8526; Colak *et al.* (1995) *Brain Res.* 691:76-82; Crystal (1995) *Science* 270:404-410; Elshami *et al.* (1996) *Human Gene Ther.* 7:141-148; Vincent *et al.* (1996) *J. Neurosurg.* 85:648-654), and many other diseases. Replication-defective retroviral vectors harboring therapeutic polynucleotide sequence as part of the retroviral genome have also been used, particularly with regard to simple MuLV vectors. See, e.g., Miller *et al.* (1990) *Mol. Cell. Biol.* 10:4239 (1990);

Kolberg (1992) *J. NIH Res.* 4:43, and Cornetta *et al.* (1991) *Hum. Gene Ther.* 2:215). Nucleic acid transport coupled to ligand-specific, cation-based transport systems (Wu and Wu (1988) *J. Biol. Chem.* 263:14621-14624) have also been used. Naked DNA expression vectors have also been described (Nabel *et al.* (1990), *supra*; Wolff *et al.* 5 (1990) *Science* 247:1465-1468). In general, these approaches can be adapted to the invention by incorporating nucleic acids encoding the interferon homologues herein into the appropriate vectors.

General texts which describe gene therapy protocols, which can be adapted to the present invention by introducing the nucleic acids of the invention into patients, 10 include Robbins (1996) *Gene Therapy Protocols*, Humana Press, NJ, and Joyner (1993) *Gene Targeting: A Practical Approach*, IRL Press, Oxford, England.

#### Antisense Technology

In addition to expression of the nucleic acids of the invention as gene replacement nucleic acids, the nucleic acids are also useful for sense and anti-sense 15 suppression of expression, *e.g.*, to down-regulate expression of a nucleic acid of the invention, once expression of the nucleic acid is no-longer desired in the cell. Similarly, the nucleic acids of the invention, or subsequences or anti-sense sequences thereof, can also be used to block expression of naturally occurring homologous nucleic acids. A variety of sense and anti-sense technologies are known in the art, *e.g.*, as set forth in 20 Lichtenstein and Nellen (1997) *Antisense Technology: A Practical Approach* IRL Press at Oxford University, Oxford, England, and in Agrawal (1996) *Antisense Therapeutics* Humana Press, NJ, and the references cited therein.

#### Pharmaceutical Compositions

The polynucleotides and polypeptides of the invention (including vectors, 25 cells, antibodies, *etc.*, comprising polynucleotides or polypeptides of the invention) may be employed for therapeutic and prophylactic uses in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically or prophylactically effective amount of the polynucleotide or polypeptide of the invention, and a pharmaceutically acceptable carrier or excipient. A pharmaceutically acceptable carrier 30 encompasses any of the standard pharmaceutical carriers, buffers and excipients. Such a carrier or excipient includes, but is not limited to, saline, buffered saline (*e.g.*, phosphate-buffered saline solution), dextrose, water, glycerol, ethanol, emulsions (such as an

oil/water or water/oil emulsion), various types of wetting agents and/or adjuvants, and combinations thereof. Suitable pharmaceutical carriers and agents are described in REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Publishing Co., Easton, 19<sup>th</sup> ed. 1995). The formulation should suit the mode of administration of the active agent (*e.g.*, 5 nucleotide, polypeptide, vector, cell, *etc.*). Methods of administering nucleic acids, polypeptides, vectors, cells, antibodies, and proteins are well known in the art, and further discussed below.

Use as Probes

Also contemplated are uses of polynucleotides, also referred to herein as 10 oligonucleotides, typically having at least 12 bases, preferably at least 15, more preferably at least 20, 30, or 50 bases, which hybridize under at least highly stringent (or ultra-high stringent or ultra-ultra- high stringent conditions) conditions to an interferon homologue polynucleotide sequence described above. The polynucleotides may be used as probes, primers, sense and antisense agents, and the like, according to methods as noted *supra*.

15 SEQUENCE VARIATIONS

Silent Variations

It will be appreciated by those skilled in the art that due to the degeneracy 20 of the genetic code, a multitude of nucleic acids sequences encoding interferon homologue polypeptides of the invention may be produced, some which may bear minimal sequence homology to the nucleic acid sequences explicitly disclosed herein.

Table 1  
Codon Table

Amino acids			Codon					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

5           For instance, inspection of the codon table (Table 1) shows that codons AGA, AGG, CGA, CGC, CGG, and CGU all encode the amino acid arginine. Thus, at every position in the nucleic acids of the invention where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described above without altering the encoded polypeptide. It is understood that U in an RNA sequence  
10          corresponds to T in a DNA sequence.

Using, as an example, the nucleic acid sequence corresponding to nucleotides 1-15 of SEQ ID NO:1, TGT GAT CTG CCT CAG, a silent variation of this sequence includes TGC GAC TTA CCA CAA, both sequences which encode the amino acid sequence CDLPQ, corresponding to amino acids 1-5 of SEQ ID NO:36.

15          Such "silent variations" are one species of "conservatively modified variations," discussed below. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified by standard techniques to encode a functionally identical polypeptide. Accordingly, each

silent variation of a nucleic acid which encodes a polypeptide is implicit in any described sequence. The invention provides each and every possible variation of nucleic acid sequence encoding a polypeptide of the invention that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code (*e.g.*, as set forth in Table 1) as applied to the nucleic acid sequence encoding an interferon homologue polypeptide of the invention. All such variations of every nucleic acid herein are specifically provided and described by consideration of the sequence in combination with the genetic code.

Conservative Variations

“Conservatively modified variations” or, simply, “conservative variations” of a particular nucleic acid sequence refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or, where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. One of skill will recognize that individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 4%, 3%, 2% or 1%) in an encoded sequence are “conservatively modified variations” where the alterations result in the deletion of an amino acid, addition of an amino acid, or substitution of an amino acid with a chemically similar amino acid.

Conservative substitution tables providing functionally similar amino acids are well known in the art. Table 2 sets forth six groups which contain amino acids that are “conservative substitutions” for one another.

**Table 2**  
Conservative Substitution Groups

1	Alanine (A)	Serine (S)	Threonine (T)
2	Aspartic acid (D)	Glutamic acid (E)	
3	Asparagine (N)	Glutamine (Q)	
4	Arginine (R)	Lysine (K)	
5	Isoleucine (I)	Leucine (L)	Methionine (M)      Valine (V)
6	Phenylalanine (F)	Tyrosine (Y)	Tryptophan (W)

25

Thus, “conservatively substituted variations” or “conservative substitutions” of a listed polypeptide sequence of the present invention include

substitutions of a small percentage, typically less than 5%, more typically less than 4%, 3%, 2% or 1%, of the amino acids of the polypeptide sequence, with a conservatively selected amino acid of the same conservative substitution group.

For example, a conservatively substituted variation of the polypeptide  
5 identified herein as SEQ ID NO:36 will contain “conservative substitutions”, according to the six groups defined above, in up to about 8 or 9 residues (*i.e.*, about 5% of the amino acids) in the 166-amino acid polypeptide.

In a further example, if four conservative substitutions were localized in  
the region corresponding to amino acid residues 141-166 of SEQ ID NO:36, examples of  
10 conservatively substituted variations of this region,

WEVVR AEIMR SFSFS TNLQK RLRRKE include:  
WEVVR SEIMR SFSYS TNLQR RLRRKD and  
WELVR AEIVR SFSFS TNLNK RLRKKE and the like, in accordance with the  
conservative substitutions listed in Table 2 (in the above example, conservative  
15 substitutions are underlined). Listing of a protein sequence herein, in conjunction with the  
above substitution table, provides an express listing of all conservatively substituted  
proteins.

Finally, the addition of sequences which do not alter the encoded activity of  
a nucleic acid molecule, such as the addition of a non-functional sequence, is a  
20 conservative variation of the basic nucleic acid.

One of ordinary skill will appreciate that many conservative variations of  
the nucleic acid constructs which are disclosed yield a functionally identical construct.  
For example, as discussed above, owing to the degeneracy of the genetic code, “silent  
substitutions” (*i.e.*, substitutions in a nucleic acid sequence which do not result in an  
25 alteration in an encoded polypeptide) are an implied feature of *every* nucleic acid sequence  
which encodes an amino acid. Similarly, “conservative amino acid substitutions,” in one  
or a few amino acids in an amino acid sequence are substituted with different amino acids  
with highly similar properties, are also readily identified as being highly similar to a  
disclosed construct. Such conservative variations of each disclosed sequence are a feature  
30 of the present invention.

Nucleic Acid Hybridization

Nucleic acids "hybridize" when they associate, typically in solution.

Nucleic acids hybridize due to a variety of well characterized physico-chemical forces, such as hydrogen bonding, solvent exclusion, base stacking and the like. An extensive 5 guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, part I, chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays," (Elsevier, New York), as well as in Ausubel, *supra*, Hames and Higgins (1995) *Gene Probes 1*, IRL Press at Oxford University Press, Oxford, 10 England (Hames and Higgins 1) and Hames and Higgins (1995) *Gene Probes 2*, IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 2) provide details on the synthesis, labeling, detection and quantification of DNA and RNA, including oligonucleotides.

"Stringent hybridization wash conditions" in the context of nucleic acid 15 hybridization experiments, such as Southern and northern hybridizations, are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993), *supra*, and in Hames and Higgins 1 and Hames and Higgins 2, *supra*.

For purposes of the present invention, generally, "highly stringent" 20 hybridization and wash conditions are selected to be about 5° C or less lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH (as noted below, highly stringent conditions can also be referred to in comparative terms). The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the test sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected 25 to be equal to the  $T_m$  for a particular probe.

The  $T_m$  is the temperature of the nucleic acid duplexes indicates the 30 temperature at which the duplex is 50% denatured under the given conditions and its represents a direct measure of the stability of the nucleic acid hybrid. Thus, the  $T_m$  corresponds to the temperature corresponding to the midpoint in transition from helix to random coil; it depends on length, nucleotide composition, and ionic strength for long stretches of nucleotides.

After hybridization, unhybridized nucleic acid material can be removed by a series of washes, the stringency of which can be adjusted depending upon the desired results. Low stringency washing conditions (e.g., using higher salt and lower temperature) increase sensitivity, but can produce nonspecific hybridization signals and high background signals. Higher stringency conditions (e.g., using lower salt and higher temperature that is closer to the hybridization temperature) lowers the background signal, typically with only the specific signal remaining. See Rapley, R. and Walker, J.M. eds., *Molecular Biotechnology Handbook* (Humana Press, Inc. 1998) (hereinafter "Rapley and Walker"), which is incorporated herein by reference in its entirety for all purposes.

10 The  $T_m$  of a DNA-DNA duplex can be estimated using the following equation:

$$T_m (\text{°C}) = 81.5 \text{°C} + 16.6 (\log_{10} M) + 0.41 (\%G + C) - 0.72 (\%f) - 500/n,$$

where M is the molarity of the monovalent cations (usually Na+), (%G +

C) is the percentage of guanosine (G) and cytosine (C) nucleotides, (%f) is the

15 percentage of formaldehyde and n is the number of nucleotide bases (*i.e.*, length) of the hybrid. See Rapley and Walker, *supra*.

The  $T_m$  of an RNA-DNA duplex can be estimated as follows:

$$T_m (\text{°C}) = 79.8 \text{°C} + 18.5 (\log_{10} M) + 0.58 (\%G + C) - 11.8(\%G + C)^2 - 0.56$$

(%f) - 820/n, where M is the molarity of the monovalent cations (usually Na+), (%G +

20 C) is the percentage of guanosine (G) and cytosine (C) nucleotides, (%f) is the percentage of formamide and n is the number of nucleotide bases (*i.e.*, length) of the hybrid. *Id.*

Equations 1 and 2 are typically accurate only for hybrid duplexes longer than about 100-200 nucleotides. *Id.*

25 The  $T_m$  of nucleic acid sequences shorter than 50 nucleotides can be calculated as follows:

$$T_m (\text{°C}) = 4(G + C) + 2(A + T),$$

where A (adenine), C, T (thymine), and G are the numbers of the corresponding nucleotides.

30 An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42°C, with

the hybridization being carried out overnight. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see Sambrook, supra* for a description of SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove background probe signal. An example low stringency wash is 2x SSC at 40°C for 15

5 minutes.

In general, a signal to noise ratio of 2.5x-5x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Detection of at least stringent hybridization between two sequences in the context of the present invention indicates relatively strong structural similarity or

10 homology to, *e.g.*, the nucleic acids of the present invention provided in the sequence listings herein.

As noted, "highly stringent" conditions are selected to be about 5° C or less lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. Target sequences that are closely related or identical to the nucleotide

15 sequence of interest (*e.g.*, "probe") can be identified under highly stringency conditions. Lower stringency conditions are appropriate for sequences that are less complementary.

*See, e.g., Rapley and Walker, supra.*

Comparative hybridization can be used to identify nucleic acids of the invention, and this comparative hybridization method is a preferred method of

20 distinguishing nucleic acids of the invention. Detection of highly stringent hybridization between two nucleotide sequences in the context of the present invention indicates relatively strong structural similarity/homology to, *e.g.*, the nucleic acids provided in the sequence listing herein. Highly stringent hybridization between two nucleotide sequences demonstrates a degree of similarity or homology of structure, nucleotide base composition,

25 arrangement or order that is greater than that detected by stringent hybridization conditions. In particular, detection of highly stringent hybridization in the context of the present invention indicates strong structural similarity or structural homology (*e.g.*, nucleotide structure, base composition, arrangement or order) to, *e.g.*, the nucleic acids provided in the sequence listings herein. For example, it is desirable to identify test

30 nucleic acids which hybridize to the exemplar nucleic acids herein under stringent conditions.

Thus, one measure of stringent hybridization is the ability to hybridize to one of the listed nucleic acids (e.g., nucleic acid sequences SEQ ID NO:1 to SEQ ID NO:35, and SEQ ID NO:72 to SEQ ID NO:78, and complementary polynucleotide sequences thereof) under highly stringent conditions (or very stringent conditions, or ultra-high stringency hybridization conditions, or ultra-ultra high stringency hybridization conditions). Stringent hybridization (including, e.g., highly stringent, ultra-high stringency, or ultra-ultra high stringency hybridization conditions) and wash conditions can easily be determined empirically for any test nucleic acid.

For example, in determining highly stringent hybridization and wash conditions, the hybridization and wash conditions are gradually increased (e.g., by increasing temperature, decreasing salt concentration, increasing detergent concentration and/or increasing the concentration of organic solvents, such as formalin, in the hybridization or wash), until a selected set of criteria are met. For example, the hybridization and wash conditions are gradually increased until a probe comprising one or more nucleic acid sequences selected from SEQ ID NO:1 to SEQ ID NO:35, SEQ ID NO:72 to SEQ ID NO:78, and complementary polynucleotide sequences thereof, binds to a perfectly matched complementary target (again, a nucleic acid comprising one or more nucleic acid sequences selected from SEQ ID NO:1 to SEQ ID NO:35, SEQ ID NO:72 to SEQ ID NO:78, and complementary polynucleotide sequences thereof), with a signal to noise ratio that is at least 2.5x, and optionally 5x or more as high as that observed for hybridization of the probe to an unmatched target. In this case, the unmatched target is a nucleic acid corresponding to a known alpha interferon, e.g., an alpha interferon nucleic acid that is present in a public database such as GenBank™ at the time of filing of the subject application. Examples of such unmatched target nucleic acids include, e.g., those with the following GenBank accession numbers: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1). Additional such sequences can be identified in GenBank by one of ordinary skill in the art. Nomenclature of the human interferon genes and proteins is discussed in Diaz et al., (1996) *J. Interferon and Cytokine Res.* 16:179-180 and Allen et al. (1996) *J.*

*Interferon and Cytokine Res.* 16:181-184, respectively, each of which is incorporated herein by reference in its entirety for all purposes.

A test nucleic acid is said to specifically hybridize to a probe nucleic acid when it hybridizes at least  $\frac{1}{2}$  as well to the probe as to the perfectly matched complementary target, i.e., with a signal to noise ratio at least  $\frac{1}{2}$  as high as hybridization of the probe to the target under conditions in which the perfectly matched probe binds to the perfectly matched complementary target with a signal to noise ratio that is at least about 2.5x-10x, typically 5x-10x as high as that observed for hybridization to any of the unmatched target nucleic acids represented by GenBank accession numbers J00210  
5 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540  
10 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289,  
V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar interferon-alpha sequences  
15 presented in GenBank.

Ultra high-stringency hybridization and wash conditions are those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched complementary target nucleic acid is at least 10x as high as that observed for hybridization to any of the unmatched target  
20 nucleic acids represented by GenBank accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955  
25 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar IFN-alpha sequences presented in GenBank A target  
nucleic acid which hybridizes to a probe under such conditions, with a signal to noise ratio of at least  $\frac{1}{2}$  that of the perfectly matched complementary target nucleic acid is said to bind to the probe under ultra-high stringency conditions.

Similarly, even higher levels of stringency can be determined by gradually  
30 increasing the hybridization and/or wash conditions of the relevant hybridization assay. For example, those in which the stringency of hybridization and wash conditions are increased until the signal to noise ratio for binding of the probe to the perfectly matched

complementary target nucleic acid is at least 10X, 20X, 50X, 100X, or 500X or more as high as that observed for hybridization to any of the unmatched target nucleic acids represented by GenBank accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 5 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 10 (alpha-Con1), or other similar interferon-alpha sequences presented in GenBank, can be identified. A target nucleic acid which hybridizes to a probe under such conditions, with a signal to noise ratio of at least  $\frac{1}{2}$  that of the perfectly matched complementary target 15 nucleic acid is said to bind to the probe under ultra-ultra-high stringency conditions.

Target nucleic acids which hybridize to the nucleic acids represented by SEQ ID NO:1 to SEQ ID NO:35 and SEQ ID NO:72 to SEQ ID NO:78 under high, ultra-high and ultra-ultra high stringency conditions are a feature of the invention. Examples of 15 such nucleic acids include those with one or a few silent or conservative nucleic acid substitutions as compared to a given nucleic acid sequence.

Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, *e.g.*, when a copy of a nucleic acid is created using the 20 maximum codon degeneracy permitted by the genetic code, or when antisera or antiserum generated against one or more of SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID NO:79 to SEQ ID NO:85 which has been subtracted using the polypeptides encoded by known 25 interferon-alpha sequences, including, *e.g.*, the those encoded by the following interferon-alpha nucleic acid sequences in GenBank: Accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar interferon-alpha sequences presented in GenBank. 30 Further details on immunological identification of polypeptides of the invention are found below. Additionally, for distinguishing between duplexes with sequences of less than about 100 nucleotides, a TMAC1 hybridization procedure known to those of ordinary skill

in the art can be used. *See, e.g., Sorg, U. et al. 1 Nucleic Acids Res.* (Sept. 11, 1991) 19(17), incorporated herein by reference in its entirety for all purposes.

In one aspect, the invention provides a nucleic acid which comprises a unique subsequence in a nucleic acid selected from SEQ ID NO:1 to SEQ ID NO:35 or 5 SEQ ID NO:72 to SEQ ID NO:78. The unique subsequence is unique as compared to a nucleic acid corresponding to any known interferon-alpha nucleic acid sequence including, *e.g.*, the known sequences represented by GenBank accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 10 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), A12109, R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar interferon-alpha 15 sequences presented in GenBank. Such unique subsequences can be determined by aligning any of SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78 against the complete set of nucleic acids corresponding to GenBank accession numbers of known interferon-alpha nucleic acid sequences, such as, *e.g.*, J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 20 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), A12109 (alpha-4B), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar interferon-alpha 25 sequences presented in GenBank. Alignment can be performed using the BLAST algorithm set to default parameters. Any unique subsequence is useful, *e.g.*, as a probe to identify the nucleic acids of the invention.

Similarly, the invention includes a polypeptide which comprises a unique amino acid subsequence in a polypeptide selected from: SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85. Here, the unique subsequence is unique as compared to an amino acid subsequence of a known interferon-alpha polypeptide including, *e.g.*, an amino acid subsequence of a polypeptide encoded by a known 30 interferon-alpha nucleic acid corresponding to any of GenBank accession numbers: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16),

V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or other similar interferon-alpha nucleic acid or polypeptide sequences presented in GenBank. Here again, the polypeptide  
5 is aligned against the complete set of known interferon-alpha polypeptide sequences, such as those polypeptides encoded by nucleic acids corresponding to GenBank accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7),  
10 X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), - M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), (referred to as "the control polypeptides") (note that where the sequence corresponds to a non-translated sequence such as a pseudo gene, the corresponding polypeptide is generated simply by in silico translation of the nucleic acid sequence into an amino acid sequence, where the reading  
15 frame is selected to correspond to the reading frame of homologous alpha interferon nucleic acids) or other similar interferon-alpha nucleic acid or polypeptide sequences presented in GenBank.

In addition, the present invention provides a target nucleic acid which hybridizes under at least stringent or highly stringent conditions (or conditions of greater stringency) to a unique coding oligonucleotide which encodes a unique subsequence in a polypeptide selected from: SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, wherein the unique subsequence is unique as compared to a an amino acid subsequence of a known interferon-alpha polypeptide sequence shown in GenBank or to a polypeptide corresponding to any of the control polypeptides. Unique sequences are determined as noted above.  
25

In one example, the stringent conditions are selected such that a perfectly complementary oligonucleotide to the coding oligonucleotide hybridizes to the coding oligonucleotide with at least about a 5-10x higher signal to noise ratio than for hybridization of the perfectly complementary oligonucleotide to a control nucleic acid  
30 corresponding to any of the control polypeptides. Conditions can be selected such that higher ratios of signal to noise are observed in the particular assay which is used, *e.g.*, about 15x, 20x, 30x, 50x or more. In this example, the target nucleic acid hybridizes to

the unique coding oligonucleotide with at least a 2x higher signal to noise ratio as compared to hybridization of the control nucleic acid to the coding oligonucleotide. Again, higher signal to noise ratios can be selected, *e.g.*, about 2.5x, about 5x, about 10x, about 20x, about 30x, about 50x or more. The particular signal will depend on the label used in the relevant assay, *e.g.*, a fluorescent label, a colorimetric label, a radio active label, or the like.

In another aspect, the invention provides a polypeptide that comprises unique subsequence in a polypeptide selected from SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID NO:79 to SEQ ID NO:85, wherein the unique subsequence is unique as compared 10 to a polypeptide sequence corresponding to a known interferon-alpha polypeptide, such as, *e.g.*, an interferon-alpha polypeptide sequence present in GenBank.

#### SUBSTRATES AND FORMATS FOR SEQUENCE RECOMBINATION

The polynucleotides of the invention are useful as substrates for a variety of recombination and recursive recombination (*e.g.*, DNA shuffling) reactions, as well as 15 other diversity generating techniques, including mutagenesis techniques and standard cloning methods as set forth in, *e.g.*, Ausubel, Berger and Sambrook, *supra*, *i.e.*, to produce additional interferon-alpha homologues with desired properties. Based on the screening or selection protocols employed, recombinant, *e.g.*, shuffled, interferon-alpha homologue polypeptides can be generated and isolated that confer a variety of desirable 20 characteristics, *e.g.*, enhanced antiviral activity, enhanced antiproliferative activity, increased growth inhibitory, cytostatic and/or cytotoxic activities towards particular target cells, reduced immunogenicity, *etc.*

A variety of diversity generating protocols, including nucleic acid shuffling protocols, are available and fully described in the art. The procedures can be used 25 separately, and/or in combination to produce one or more variants of a nucleic acid or set of nucleic acids, as well variants of encoded proteins. Individually and collectively, these procedures provide robust, widely applicable ways of generating diversified nucleic acids and sets of nucleic acids (including, *e.g.*, nucleic acid libraries) useful, *e.g.*, for the engineering or rapid evolution of nucleic acids, proteins, pathways, cells and/or organisms 30 with new and/or improved characteristics.

While distinctions and classifications are made in the course of the ensuing discussion for clarity, it will be appreciated that the techniques are often not mutually

exclusive. Indeed, the various methods can be used singly or in combination, in parallel or in series, to access diverse sequence variants.

The result of any of the diversity generating procedures described herein can be the generation of one or more nucleic acids, which can be selected or screened for 5 nucleic acids that encode proteins with or which confer desirable properties. Following diversification by one or more of the methods herein, or otherwise available to one of skill, any nucleic acids that are produced can be selected for a desired activity or property, *e.g.*, enhanced antiviral activity, enhanced antiproliferative activity, enhanced anti-angiogenic activity, increased growth inhibitory, cytostatic and/or cytotoxic activities towards 10 particular target cells, reduced immunogenicity, *etc.* Methods for determining nucleic acids having enhanced antiviral, antiproliferative, growth inhibitory, cytostatic, and/or cytotoxic activity or reduced immunogenicity include those described herein. This can include identifying any activity that can be detected, for example, in an automated or automatable format, by any of the assays in the art. A variety of related (or even 15 unrelated) properties can be evaluated, in serial or in parallel, at the discretion of the practitioner.

The following publications describe a variety of diversity generating procedures, including recursive recombination procedures, and/or methods for generating modified nucleic acid sequences for use in the procedures and methods of the present 20 invention include the following publications and the references cited therein: Soong, N. W. *et al.* (2000) "Molecular Breeding of Viruses," Nature Genetics 25:436-439; Stemmer, W. *et al.* (1999) "Molecular breeding of viruses for targeting and other clinical properties," Tumor Targeting 4:1-4; Ness *et al.* (1999) "DNA Shuffling of subgenomic sequences of subtilisin," Nature Biotechnology 17:893-896; Chang *et al.* (1999) 25 "Evolution of a cytokine using DNA family shuffling," Nature Biotechnology 17:793-797; Minshull and Stemmer (1999) "Protein evolution by molecular breeding," Current Opinion in Chemical Biology 3:284-290; Christians *et al.* (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling," Nature Biotechnology 17:259-264; Crameri *et al.* (1998) "DNA shuffling of a family of genes 30 from diverse species accelerates directed evolution," Nature 391:288-291; Crameri *et al.* (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," Nature Biotechnology 15:436-438; Zhang *et al.* (1997) "Directed evolution of an effective

fucosidase from a galactosidase by DNA shuffling and screening," *Proc. Nat'l Acad. Sci. USA* 94:4504-4509; Patten *et al.* (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines," *Current Opinion in Biotechnology* 8:724-733; Crameri *et al.* (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling," 5 *Nature Medicine* 2:100-103; Crameri *et al.* (1996) "Improved green fluorescent protein by molecular evolution using DNA shuffling," *Nature Biotechnology* 14:315-319; Gates *et al.* (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer,'" *J. Mol. Biol.* 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: *The Encyclopedia of Molecular Biology*, VCH Publishers, New York. pp. 447-457; Crameri and Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes," *BioTechniques* 18:194-195; Stemmer *et al.* (1995) "Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxy-ribonucleotides" *Gene* 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation," *Science* 270:1510; Stemmer (1995) 10 "Searching Sequence Space," *Bio/Technology* 13:549-553; Stemmer (1994) "Rapid evolution of a protein *in vitro* by DNA shuffling," *Nature* 370:389-391; and Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution," *Proc. Nat'l Acad. Sci. USA* 91:10747-10751.

Additional details regarding DNA shuffling and other diversity generating 20 methods can be found in the following U.S. patents, PCT publications, and EP publications: USPN 5,605,793 to Stemmer (February 25, 1997), "Methods for *In vitro* Recombination;" USPN 5,811,238 to Stemmer et al. (September 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" USPN 5,830,721 to Stemmer et al. (November 3, 1998), "DNA 25 Mutagenesis by Random Fragmentation and Reassembly;" USPN 5,834,252 to Stemmer (November 10, 1998) "End-Complementary Polymerase Reaction;" USPN 5,837,458 to Minshull (November 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz, "End 30 Complementary Polymerase Chain Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull and Stemmer, "Methods and

Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al.  
"Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen *et al.*, "Antigen  
Library Immunization;" WO 99/41369 by Punnonen *et al.*, "Genetic Vaccine Vector  
Engineering;" WO 99/41368 by Punnonen *et al.*, "Optimization of Immunomodulatory  
5 Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA  
Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer  
"Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107  
by Stemmer *et al.*, "Modification of Virus Tropism and Host Range by Viral Genome  
Shuffling," WO 99/21979 by Apt *et al.*, "Human Papillomavirus Vectors;" WO 98/31837  
10 by Del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence  
Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for  
Polypeptide Engineering;" EP 0946755 by Patten and Stemmer, "Methods and  
Compositions for Polypeptide Engineering;" and WO 98/13487 by Stemmer *et al.*,  
"Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and  
15 Selection;" WO 00/00632, "Methods for Generating Highly Diverse Libraries," WO  
00/09679, "Methods for Obtaining *in vitro* Recombined Polynucleotide Sequence Banks  
and Resulting Sequences," WO 98/42832 by Arnold *et al.*, "Recombination of  
Polynucleotide Sequences Using Random or Defined Primers," WO 99/29902 by Arnold  
*et al.*, "Method for Creating Polynucleotide and Polypeptide Sequences," WO 98/41653  
20 by Vind, "An *in vitro* Method for Construction of a DNA Library," WO 98/41622 by  
Borchert et al., "Method for Constructing a Library Using DNA Shuffling," and WO  
98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination."  
Certain U.S. applications provide additional details regarding DNA  
shuffling and related techniques, as well as other diversity generating methods, including  
25 "SHUFFLING OF CODON ALTERED GENES" by Patten et al. filed September 29,  
1998 (USSN 60/102,362), January 29, 1999 (USSN 60/117,729), and September 28, 1999  
(USSN 09/407,800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY  
RECURSIVE SEQUENCE RECOMBINATION", by Del Cardayre et al. filed July 15,  
1998 (USSN 09/166,188), and July 15, 1999 (USSN 09/354,922);  
30 "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by  
Crameri et al., filed February 5, 1999 (USSN 60/118,813), June 24, 1999 (USSN  
60/141,049), and September 28, 1999 (USSN 09/408,392); "USE OF CODON-BASED

- OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch et al., filed September 28, 1999 (USSN 09/408,393); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov and Stemmer, filed February 5, 1999  
5 (USSN 60/118854) and October 12, 1999 (USSN 09/416,375); RECOMBINATION OF INSERTION MODIFIED NUCLEIC ACIDS by Patten *et al.*, filed March 5, 1999 (USSN 60/122,943), July 2, 1999 (USSN 60/142,299), November 10, 1999 (USSN 60/164,618), and November 10, 1999 (USSN 60/164,617); and "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID  
10 FRAGMENT ISOLATION" by Affholter, USSN 60/186,482 filed March 2, 2000.

As a review of the foregoing publications, patents, published foreign applications and U.S. patent applications reveals, diversity generation methods, such as shuffling (or "recursive recombination") of nucleic acids, to provide new nucleic acids with desired properties can be carried out by a number of established methods. Any of  
15 these methods can be adapted to the present invention to evolve the alpha interferons discussed herein to produce new alpha interferon homologues with new or improved properties. Both the methods of making such interferons and the interferons (*e.g.*, IFN homologues) produced by these methods are a feature of the invention. In brief, several different general classes of sequence modification methods, such as recombination, are  
20 applicable to the present invention and set forth, *e.g.*, in the references above. First, nucleic acids can be recombined *in vitro* by any of a variety of techniques discussed in the references above, including *e.g.*, DNase digestion of nucleic acids to be recombined followed by ligation and/or PCR reassembly of the nucleic acids. Second, nucleic acids can be recursively recombined *in vivo* or *ex vivo*, *e.g.*, by allowing recombination to occur  
25 between nucleic acids in cells. Third, whole genome recombination methods can be used in which whole genomes of cells or other organisms are recombined, optionally including spiking of the genomic recombination mixtures with desired library components (*e.g.*, genes corresponding to the pathways of the present invention). Fourth, synthetic recombination methods can be used, in which oligonucleotides corresponding to targets of  
30 interest are synthesized and reassembled in PCR or ligation reactions which include oligonucleotides which correspond to more than one parental nucleic acid, thereby generating new recombinant nucleic acids. Oligonucleotides can be made by standard

nucleotide addition methods, or can be made, *e.g.*, by tri-nucleotide synthetic approaches. Fifth, *in silico* methods of recombination can be effected in which genetic algorithms are used in a computer to recombine sequence strings which correspond to homologous (or even non-homologous) nucleic acids. The resulting recombined sequence strings are 5 optionally converted into nucleic acids by synthesis of nucleic acids which correspond to the recombined sequences, *e.g.*, in concert with oligonucleotide synthesis/ gene reassembly techniques. Any of the preceding general recombination formats can be practiced in a reiterative fashion to generate a more diverse set of recombinant nucleic acids. Sixth, methods of accessing natural diversity, *e.g.*, by hybridization of diverse 10 nucleic acids or nucleic acid fragments to single-stranded templates, followed by polymerization and/or ligation to regenerate full-length sequences, optionally followed by degradation of the templates and recovery of the resulting modified nucleic acids can be used. above references provide these and other basic recombination formats as well as many modifications of these formats. Regardless of the format which is used, the nucleic 15 acids of the invention can be recombined (with each other, or with related (or even unrelated) nucleic acids to produce a diverse set of recombinant nucleic acids, including *e.g.*, homologous nucleic acids. In general, the sequence recombination techniques described herein provide particular advantages in that they provide for recombination between the nucleic acids of SEQ ID NO:1 to SEQ ID NO:35, and SEQ ID NO:72 to SEQ 20 ID NO:78, or fragments or variants thereof, in any available format, thereby providing a very fast way of exploring the manner in which different combinations of sequences can affect a desired result.

Following recombination, any nucleic acids which are produced can be screened or selected for a desired activity. In the context of the present invention, this can 25 include testing for and identifying any activity that can be detected, *e.g.*, in an automatable format, by any assay known in the art. In addition, useful properties such as low immunogenicity, increased half-life, improved solubility, oral availability, or the like can also be selected for. A variety of alpha-interferon related (or even unrelated) properties can be assayed for, using any available assay.

30 DNA mutagenesis and shuffling provide a robust, widely applicable, means of generating diversity useful for the engineering of proteins, pathways, cells and organisms with improved characteristics. In addition to the basic formats described above,

it is sometimes desirable to combine shuffling methodologies with other techniques for generating diversity. In conjunction with (or separately from) shuffling methods, a variety of diversity generation methods can be practiced and the results (*i.e.*, diverse populations of nucleic acids) screened for in the systems of the invention. Additional diversity can be  
5 introduced by methods which result in the alteration of individual nucleotides or groups of contiguous or non-contiguous nucleotides, *i.e.*, mutagenesis methods. Many mutagenesis methods are found in the above-cited references; additional details regarding mutagenesis methods can be found in the references listed below.

Mutagenesis methods of generating diversity include, for example,  
10 recombination (PCT/US98/05223; Publ. No. WO98/42727); site-directed mutagenesis (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview," Anal. Biochem. 254(2):157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method," Methods Mol. Biol. 57:369-374; Smith (1985) "*In vitro* mutagenesis," Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and  
15 applications of *in vitro* mutagenesis," Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis," Biochem. J. 237:1-7; and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis," in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without  
20 phenotypic selection," Proc. Nat'l Acad. Sci. USA 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection," Results Probl. Cell Differ. 154, 367-382; and Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities," Science 242:240-245); oligonucleotide-directed mutagenesis (Results Probl. Cell Differ. 100:468-500 (1983); Results Probl. Cell Differ. 154:329-350  
25 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment," Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983)  
"Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors,"  
Results Probl. Cell Differ. 100:468-500; and Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template," Results Probl. Cell Differ. 154:329-350); phosphorothioate-modified DNA mutagenesis (Taylor et al. (1985) "The use of phosphorothioate-modified

DNA in restriction enzyme reactions to prepare nicked DNA," Nucl. Acids Res. 13:8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA," Nucl. Acids Res. 13:8765-8787 (1985); Nakamaye & Eckstein (1986) "Inhibition of restriction endonuclease Nci I 5 cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis," Nucl. Acids Res. 14:9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis," Nucl. Acids Res. 16:791-802; and Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide," 10 Nucl. Acids Res. 16:803-814); mutagenesis using gapped duplex DNA (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction," Nucl. Acids Res. 12:9441-9456; Kramer & Fritz (1987) "Oligonucleotide-directed construction of mutations via gapped duplex DNA," Results Probl. Cell Differ. 15 154:350-367; Kramer et al. (1988) "Improved enzymatic *in vitro* reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations," Nucl. Acids Res. 16:7207; and Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions *in vitro*," Nucl. Acids Res. 16:6987-6999).

Additional suitable methods include point mismatch repair (Kramer et al. 20 (1984) "Point Mismatch Repair," Cell 38:879-887), mutagenesis using repair-deficient host strains (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors," Nucl. Acids Res. 13:4431-4443; and Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors," Results Probl. Cell Differ. 154:382-403), deletion mutagenesis (Eghitedarzadeh & Henikoff (1986) "Use of 25 oligonucleotides to generate large deletions," Nucl. Acids Res. 14:5115), restriction-selection and restriction-selection and restriction-purification (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin," Phil. Trans. R. Soc. Lond. A 317:415-423), mutagenesis by total gene synthesis (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein," 30 Science 223:1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducing)," Nucl. Acids Res. 14:6361-6372; Wells et al. (1985) "Cassette mutagenesis:

an efficient method for generation of multiple mutations at defined sites," *Gene* 34:315-323; and Grundström et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis," *Nucl. Acids Res.* 13:3305-3316), double-strand break repair (Mandecki (1986) "Oligonucleotide-directed double-strand break repair in plasmids of 5 *Escherichia coli*: a method for site-specific mutagenesis," *Proc. Nat'l Acad. Sci. USA*, 83:7177-7181). Additional details on many of the above methods can be found in *Methods in Enzymology*, Vol. 154, which also describes useful controls for trouble-shooting problems with various mutagenesis methods.

Random or semi-random mutagenesis using doped or degenerate 10 oligonucleotides (Arkin and Youvan (1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis," *Biotechnology* 10:297-300; Reidhaar-Olson et al. (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes," *Methods Enzymol.* 208:564-86; Lim and Sauer (1991) "The role of internal packing interactions in determining the structure and stability of a protein," 15 *J. Mol. Biol.* 219:359-76; Breyer and Sauer (1989) "Mutational analysis of the fine specificity of binding of monoclonal antibody 51F to lambda repressor," *J. Biol. Chem.* 264:13355-60); "Walk-Through Mutagenesis" (Crea, R.; US Patents 5,830,650 and 5,798,208, and EP Patent 0527809 B1) may also be employed to generate diversity.

In one aspect of the present invention, error-prone PCR can be used to 20 generate nucleic acid variants. Using this technique, PCR is performed under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Examples of such techniques are found in the references above and, e.g., in Leung et al. (1989) *Technique* 1:11-15 and Caldwell et al. (1992) *PCR Methods Applic.* 2:28-33. Similarly, assembly 25 PCR can be used, in a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions can occur in parallel in the same vial, with the products of one reaction priming the products of another reaction. Sexual PCR mutagenesis can be used in which homologous recombination occurs between DNA molecules of different but related DNA sequence *in vitro*, by 30 random fragmentation of the DNA molecule based on sequence homology, followed by fixation of the crossover by primer extension in a PCR reaction. This process is described in the references above, e.g., in Stemmer (1994) *Proc. Nat'l Acad. Sci. USA* 91:10747-

10751. Recursive ensemble mutagenesis can be used in which an algorithm for protein mutagenesis is used to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. Examples of this  
5 approach are found in Arkin & Youvan (1992) Proc. Nat'l Acad. Sci. USA 89:7811-7815.

As noted, oligonucleotide directed mutagenesis can be used in a process which allows for the generation of site-specific mutations in any nucleic acid sequence of interest. Examples of such techniques are found in the references above and, e.g., in Reidhaar-Olson et al. (1988) Science, 241:53-57. Similarly, cassette mutagenesis can be  
10 used in a process which replaces a small region of a double stranded DNA molecule with a synthetic oligonucleotide cassette that differs from the native sequence. The oligonucleotide can contain, e.g., completely and/or partially randomized native sequence(s).

*In vivo* (or *ex vivo*) mutagenesis can be used in a process of generating  
15 random mutations in any cloned DNA of interest which involves the propagation of the DNA, e.g., in a strain of *E. coli* that carries mutations in one or more of the DNA repair pathways. These "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in one of these strains will eventually generate random mutations within the DNA.

20 Exponential ensemble mutagenesis can be used for generating combinatorial libraries with a high percentage of unique and functional mutants, where small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Examples of such procedures are found in Delegrave & Youvan (1993) Biotechnology Research 11:1548-1552. Similarly, random  
25 and site-directed mutagenesis can be used. Examples of such procedures are found in Arnold (1993) Current Opinion in Biotechnology 4:450-455.

Kits for mutagenesis, library construction, and other diversity generation methods are also commercially available. For example, kits are available from, e.g., Stratagene (e.g., QuickChange™ site-directed mutagenesis kit; and Chameleon™ double-stranded, site-directed mutagenesis kit), Bio/Can Scientific, Bio-Rad (e.g., using the Kunkel method described above), Boehringer Mannheim Corp., Clonetech Laboratories, DNA Technologies, Epicentre Technologies (e.g., 5 prime 3 prime kit); Genpak Inc,

Lemargo Inc, Life Technologies (Gibco BRL), New England Biolabs, Pharmacia Biotech, Promega Corp., Quantum Biotechnologies, Amersham International plc (*e.g.*, using the Eckstein method above), and Anglian Biotechnology Ltd (*e.g.*, using the Carter/Winter method above).

- 5 Any of the described shuffling or mutagenesis techniques can be used in conjunction with procedures which introduce additional diversity into a genome, *e.g.*, a bacterial, fungal, animal or plant genome. For example, in addition to the methods above, techniques have been proposed which produce chimeric nucleic acid multimers suitable for transformation into a variety of species (*see, e.g.*, Schellenberger U.S. Patent No. 10 5,756,316 and the references above). When such chimeric multimers consist of genes that are divergent with respect to one another (*e.g.*, derived from natural diversity or through application of site directed mutagenesis, error prone PCR, passage through mutagenic bacterial strains, and the like), are transformed into a suitable host, this provides a source of nucleic acid diversity for DNA diversification.
- 15 Chimeric multimers transformed into host species are suitable as substrates for *in vivo* (or *ex vivo*) shuffling protocols. Alternatively, a multiplicity of polynucleotides sharing regions of partial sequence similarity or homology can be transformed into a host species and recombined *in vivo* (or *ex vivo*) by the host cell. Subsequent rounds of cell division can be used to generate libraries, members of which, comprise a single, 20 homogenous population of monomeric or pooled nucleic acid. Alternatively, the monomeric nucleic acid can be recovered by standard techniques and recursively recombined in any of the described shuffling formats.

Chain termination methods of diversity generation have also been proposed (see, *e.g.*, U.S. Patent No. 5,965,408 and the references above). In this approach, double 25 stranded DNAs corresponding to one or more genes sharing regions of sequence similarity or homology are combined and denature, in the presence or absence of primers specific for the gene. The single stranded polynucleotides are then annealed and incubated in the presence of a polymerase and a chain terminating reagent (*e.g.*, ultraviolet, gamma or X-ray irradiation; ethidium bromide or other intercalators; DNA binding proteins, such as 30 single strand binding proteins, transcription activating factors, or histones; polycyclic aromatic hydrocarbons; trivalent chromium or a trivalent chromium salt; or abbreviated polymerization mediated by rapid thermocycling; and the like), resulting in the production

of partial duplex molecules. The partial duplex molecules, *e.g.*, containing partially extended chains, are then denatured and reannealed in subsequent rounds of replication or partial replication resulting in polynucleotides which share varying degrees of sequence similarity or homology and which are chimeric with respect to the starting population of 5 DNA molecules. Optionally, the products or partial pools of the products can be amplified at one or more stages in the process. Polynucleotides produced by a chain termination method, such as described above are suitable substrates for diversity generation methods (*e.g.*, RSR, DNA shuffling) according to any of the described formats.

Diversity can be further increased by using methods which are not 10 homology based with DNA shuffling (which, as set forth in the above publications and applications can be homology or non-homology based, depending on the precise format). For example, incremental truncation for the creation of hybrid enzymes (ITCHY) described in Ostermeier et al. (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology" *Nature Biotech.* 17:1205, can be used to generate an 15 initial recombinant library which serves as a substrate for one or more rounds of *in vitro*, *ex vivo*, or *in vivo* diversity generation methods (*e.g.*, RSR or shuffling methods).

Methods for generating multispecies expression libraries have been 20 described (*e.g.*, U.S. Patent Nos. 5,783,431; 5,824,485 and the references above) and their use to identify protein activities of interest has been proposed (U.S. Patent 5,958,672 and the references above). Multispecies expression libraries are, in general, libraries comprising cDNA or genomic sequences from a plurality of species or strains, operably linked to appropriate regulatory sequences, in an expression cassette. The cDNA and/or genomic sequences are optionally randomly concatenated to further enhance diversity. The vector can be a shuttle vector suitable for transformation and expression in more than 25 one species of host organism, *e.g.*, bacterial species, eukaryotic cells. In some cases, the library is biased by preselecting sequences which encode a protein of interest, or which hybridize to a nucleic acid of interest. Any such libraries can be provided as substrates for any of the methods herein described.

In some applications, it is desirable to preselect or prescreen libraries (*e.g.*, 30 an amplified library, a genomic library, a cDNA library, a normalized library, *etc.*) or other substrate nucleic acids prior to shuffling, or to otherwise bias the substrates towards nucleic acids that encode functional products (shuffling procedures can also,

independently have these effects). For example, in the case of antibody engineering, it is possible to bias the shuffling process toward antibodies with functional antigen binding sites by taking advantage of *in vivo* (or *ex vivo* or *in vitro*) recombination events prior to diversity generation (*e.g.*, DNA shuffling) by any described method. For example,

- 5 recombined CDRs derived from B cell cDNA libraries can be amplified and assembled into framework regions (*e.g.*, Jirholt *et al.* (1998) "Exploiting sequence space: shuffling *in vivo* formed complementarity determining regions into a master framework," *Gene* 215:471) prior to diversity generation (*e.g.*, DNA shuffling) according to any of the methods described herein.

10 Libraries can be biased towards nucleic acids which encode proteins with desirable activities (*e.g.*, binding affinities, enzymatic activities, anti-viral activities, ability to induce an immune response, antiproliferative activities, adjuvant properties, *etc.*). For example, after identifying a clone from a library which exhibits a specified activity, the clone can be mutagenized using any known method for introducing DNA

15 alterations, including, but not restricted to, DNA shuffling or another form of recursive sequence recombination or diversity generation. A library comprising the mutagenized homologues is then screened for a desired activity, which can be the same as or different from the initially specified activity. An example of such a procedure is proposed in U.S. Patent No. 5,939,250. Desired activities can be identified by any method known in the art.

20 For example, WO 99/10539 proposes that gene libraries can be screened by combining extracts from the gene library with components obtained from metabolically rich cells and identifying combinations which exhibit the desired activity. It has also been proposed (*e.g.*, WO 98/58085) that clones with desired activities can be identified by inserting bioactive substrates into samples of the library, and detecting bioactive fluorescence corresponding to the product of a desired activity using a fluorescent analyzer, *e.g.*, a flow 25 cytometry device, a CCD, a fluorometer, or a spectrophotometer.

Libraries can also be biased towards nucleic acids which have specified characteristics, *e.g.*, hybridization to a selected nucleic acid probe. For example, application WO 99/10539 proposes that polynucleotides encoding a desired activity (*e.g.*, 30 an enzymatic activity, for example: a lipase, an esterase, a protease, a glycosidase, a glycosyl transferase, a phosphatase, a kinase, an oxygenase, a peroxidase, a hydrolase, a hydratase, a nitrilase, a transaminase, an amidase or an acylase) can be identified from

among genomic DNA sequences in the following manner. Single stranded DNA molecules from a population of genomic DNA are hybridized to a ligand-conjugated probe. The genomic DNA can be derived from either a cultivated or uncultivated microorganism, or from an environmental sample. Alternatively, the genomic DNA can be  
5 derived from a multicellular organism, or a tissue derived therefrom.

Second strand synthesis can be conducted directly from the hybridization probe used in the capture, with or without prior release from the capture medium or by a wide variety of other strategies known in the art. Alternatively, the isolated single-stranded genomic DNA population can be fragmented without further cloning and used  
10 directly in a shuffling-based gene reassembly process. In one such method the fragment population derived the genomic library(ies) is annealed with partial, or, often approximately full length ssDNA or RNA corresponding to the opposite strand. Assembly of complex chimeric genes from this population is the mediated by nuclease-base removal of non-hybridizing fragment ends, polymerization to fill gaps between such fragments and  
15 subsequent single stranded ligation. The parental strand can be removed by digestion (if RNA or uracil-containing), magnetic separation under denaturing conditions (if labeled in a manner conducive to such separation) and other available separation/purification methods. Alternatively, the parental strand is optionally co-purified with the chimeric strands and removed during subsequent screening and processing steps. As set forth in  
20 "Single-stranded nucleic acid template-mediated recombination and nucleic acid fragment isolation" by Affholter (USSN 60/186,482, filed March 2, 2000) and WO 98/27230, "Methods and Compositions for Polypeptide Engineering" by Patten and Stemmer, shuffling using single-stranded templates and nucleic acids of interest which bind to a portion of the template can also be performed.

25 In one approach, single-stranded molecules are converted to double-stranded DNA (dsDNA) and the dsDNA molecules are bound to a solid support by ligand-mediated binding. After separation of unbound DNA, the selected DNA molecules are released from the support and introduced into a suitable host cell to generate a library enriched sequences which hybridize to the probe. A library produced in this manner  
30 provides a desirable substrate for any of the shuffling reactions described herein.

"Non-Stochastic" methods of generating nucleic acids and polypeptides are alleged in Short, J. "Non-Stochastic Generation of Genetic Vaccines and Enzymes," WO

00/46344. These methods, including the proposed non-stochastic polynucleotide reassembly and gene site saturation mutagenesis and synthetic ligation polynucleotide reassembly methods outlined therein, can be applied to the present invention as well.

It will readily be appreciated that any of the above described techniques  
5 suitable for enriching a library prior to diversification can also be used to screen the products, or libraries of products, produced by the diversity generating methods.

A recombinant nucleic acid produced by recursively recombining one or more polynucleotides of the invention with one or more additional nucleic acids also forms a part of the invention. The one or more additional nucleic acids may include  
10 another polynucleotide of the invention; optionally, alternatively, or in addition, the one or more additional nucleic acids can include, *e.g.*, a nucleic acid encoding a naturally-occurring interferon-alpha or a subsequence thereof, or any homologous interferon-alpha sequence or subsequence thereof, or an interferon-beta sequence or subsequence thereof (e.g., an interferon-alpha or interferon-beta sequence as found in GenBank or other  
15 available literature), or, *e.g.*, any other homologous or non-homologous nucleic acid (certain recombination formats noted above, notably those performed synthetically or in *silico*, do not require homology for recombination).

The recombining steps may be performed *in vivo*, *ex vivo*, *in vitro*, or *in silico* as described in more detail in the references above. Also included in the invention  
20 is a cell containing any resulting recombinant nucleic acid, nucleic acid libraries produced by diversity generation, recombination, or recursive recombination of the nucleic acids set forth herein, and populations of cells, vectors, viruses, plasmids or the like comprising the library or comprising any recombinant nucleic acid resulting from diversity generation or recombination (or recursive recombination) of a nucleic acid as set forth herein with  
25 another such nucleic acid, or an additional nucleic acid. Corresponding sequence strings in a database present in a computer system or computer readable medium are a feature of the invention.

#### OTHER POLYNUCLEOTIDE COMPOSITIONS

The invention also includes compositions comprising two or more  
30 polynucleotides of the invention (*e.g.*, as substrates for recombination). The composition can comprise a library of recombinant nucleic acids, where the library contains at least 2,

3, 5, 10, 20, or 50 or more nucleic acids. The nucleic acids are optionally cloned into expression vectors, providing expression libraries.

- The invention also includes compositions produced by digesting one or more polynucleotides of the invention with a restriction endonuclease, an RNase, or a 5 DNAse (*e.g.*, as is performed in certain of the recombination formats noted above); and compositions produced by fragmenting or shearing one or more polynucleotides of the invention by mechanical means (*e.g.*, sonication, vortexing, and the like), which can also be used to provide substrates for recombination in the methods above. Similarly, 10 compositions comprising sets of oligonucleotides corresponding to more than one nucleic acids of the invention are useful as recombination substrates and are a feature of the invention. For convenience, these fragmented, sheared, or oligonucleotide synthesized mixtures are referred to as fragmented nucleic acid sets.

- Also included in the invention are compositions produced by incubating one or more of the fragmented nucleic acid sets in the presence of ribonucleotide- or 15 deoxyribonucleotide triphosphates and a nucleic acid polymerase. This resulting composition forms a recombination mixture for many of the recombination formats noted above. The nucleic acid polymerase may be an RNA polymerase, a DNA polymerase, or an RNA-directed DNA polymerase (*e.g.*, a "reverse transcriptase"); the polymerase can be, *e.g.*, a thermostable DNA polymerase (such as, VENT, TAQ, or the like).

20 **INTERFERON HOMOLOGUE POLYPEPTIDES**

- The invention provides isolated or recombinant interferon-alpha homologue polypeptides, also referred to herein as "interferon-alpha homologues," or "interferon homologues" or "IFN-alpha homologues" or "IFN homologues". An isolated or recombinant interferon homologue polypeptide of the invention includes a polypeptide 25 comprising a sequence selected from SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID NO:79 to SEQ ID NO:85, and conservatively modified variations thereof, and fragments thereof having an antiproliferative activity in, *e.g.*, a human Daudi cell line-based assay (or other similar assay) and/or an antiviral activity in, *e.g.*, a murine cell line/EMCV-based assay (or other similar assay). An alignment of exemplary interferon homologue 30 polypeptide sequences according to the invention is provided in Fig. 1. Alignment of the polypeptide sequences of the invention to each other or to sequences of known, naturally-

occurring interferon-alphas is readily performed by one of ordinary skill in the art using publicly available databases and alignment programs.

The invention also provides a polypeptide comprising at least about 100, 120, 130, 140, 150, 155, 160, 163, 165, or 166 contiguous amino acids of any one of SEQ ID NOS:36-70 or SEQ ID NO:71. In one aspect, said amino acid sequence comprises amino acids Lys160 and Glu166, wherein the numbering of the amino acids in the sequence corresponds to that of SEQ ID NO:36.

Several conclusions may be drawn from comparison of the exemplary sequences of the invention (Fig. 1) to sequences of known, naturally-occurring interferon-alphas and other Type I interferons (including beta, delta, omega, and tau-interferons) from human and non-human sources. Such sequences are readily available from a variety of sources, such as GenBank, and the Pfam (Protein Families) database at <http://www.sanger.ac.uk/Software/Pfam/index.shtml>.

Of particular note is the presence, in some interferon homologue polypeptide sequences of the invention, of the following amino acid residues (denoted "Group I" residues) which do not appear in the equivalent position of known, naturally-occurring human or non-human Type 1 interferon sequences.

Group I: Asp11; Pro14; Arg50; Phe55; Asp75; Asn80; Pro111; Leu124; Glu134; Ser140, and Ala143; with residue numbering corresponding to the mature interferon homologue sequence identified as SEQ ID NO:36.

Also of note is the presence, in some interferon homologue polypeptide sequences of the invention, of the following amino acid residues (denoted "Group II" residues) which do not appear in the equivalent position of known, naturally-occurring human interferon-alpha subtype sequences.

Group II: Pro9; (Lys, Ser)12; (Thr, Val)24; Gln34; Arg40; Ser45; Arg47; Leu56; Ile60; Phe67; Ala79, Gly88; His90; Arg91; Glu95; Val101; (Gly, Ala)104; Val112; Gly114; Pro116; Lys133, and His136.

In other embodiments, the interferon homologue polypeptide comprises at least 20, 50, 100, 150, 155, or 160 or more contiguous amino acids of any one of SEQ ID NOS:36-70 and/or one or more of amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, wherein the numbering of the amino acids corresponds to that of SEQ ID NO:36, or one or more of

amino acids Pro9, (Lys or Ser)12, (Thr or Val)24, Gln34, Arg40, Ser45, Arg47, Leu56, Ile60, Phe67, Ala79, Gly88, His90, Arg91, Glu95, Val101, (Gly, Ala)104, Val112, Gly114, Pro116, Lys133, and His136, wherein the numbering of the amino acids in said polypeptide sequence corresponds to the numbering of individual amino acids in the  
5 amino acid sequence of SEQ ID NO:36. Thus, for example, in this embodiment, an interferon polypeptide comprises an amino acid sequence comprising a proline residue at amino acid position 9 in the sequence, a lysine or serine residue at position 12, a threonine or valine residue at position 24, a glutamine residue at position 34, an arginine residue at position 40, etc. Such polypeptides may exhibit antiproliferative activities in a human  
10 Daudi cell line-based proliferation assay (e.g., at least about  $8.3 \times 10^6$  units/mg) and/or an antiviral activities in a human WISH cell/EMCV-based assay (at least about  $2.1 \times 10^7$  units/mg). Some such polypeptides bind a human alpha interferon receptor. Some such polypeptides are 166 amino acids in length. In another aspect, such polypeptides may comprise a sequence selected from any of the group of SEQ ID NO:36 to SEQ ID NO:54.

15 An antiproliferative activity of any polypeptide of the invention generally relates to the capability or ability of a polypeptide to cause cells or parts thereof to grow or produce new cellular growth rapidly and often repeatedly.

The invention further includes a polypeptide (e.g., any of SEQ ID NOS:36-  
71 or SEQ ID NOS:79-85) or a nucleic acid (e.g., any of SEQ ID NOS:1-35 or SEQ ID  
20 NOS:72-78)encoding a polypeptide, wherein said polypeptide having an anti-angiogenic activity as measured by an anti-angiogenesis assay well known to those of ordinary skill in the art.

The invention further includes:

(a) any interferon-alpha polypeptide comprising one or more Group I  
25 amino acid residues above.

(b) any interferon-alpha polypeptide comprising one or more Group II  
amino acid residues above in the context of a human like interferon sequence (i.e., a  
sequence which displays a high level of similarity or homology to a human interferon), or  
a sequence which is highly similar or homologous (i.e., having a percent sequence  
30 homology or sequence identity of at least about 80%, 90%, 95%, 96%, 97%, 98% or  
more) to any sequence listed in the attached sequence listing or fragment thereof.

(c) any interferon-alpha polypeptide containing a combination of the following residues, which are localized in or near the regions of the interferon-alpha molecule known or proposed to interact with a Type I interferon receptor, where such sequence combinations (motifs) do not appear in the equivalent position of any known naturally-occurring human or non-human Type 1 interferon:

- 5 (i) (Tyr or Gln)34; plus one or more of Ile132 or Arg134; or
- (ii) Asp78, Glu79, or (Asp or Thr)80; plus one or more of Ile132 or Arg134.

In another embodiment, the present invention provides an interferon alpha homologue comprising the sequence show in SEQ ID NO:71: CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QQLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof, where X<sub>11</sub> is N or D; X<sub>12</sub> is R, S, or K; X<sub>15</sub> is L or M; X<sub>16</sub> is I, M, or V; X<sub>19</sub> is A or G; X<sub>22</sub> is G or R; X<sub>24</sub> is I or T; X<sub>26</sub> is P or H; X<sub>34</sub> is H, Y or Q; X<sub>38</sub> is F or L; X<sub>40</sub> is Q or R; X<sub>45</sub> is G or S; X<sub>46</sub> is N or H; X<sub>47</sub> is Q or R; X<sub>50</sub> is K or R; X<sub>51</sub> is A or T; X<sub>55</sub> is S or F; X<sub>56</sub> is V or A; X<sub>57</sub> is L or F; X<sub>60</sub> is M or I; X<sub>61</sub> is I or M; X<sub>67</sub> is L or F; X<sub>72</sub> is D or N; X<sub>75</sub> is A or V; X<sub>76</sub> is A or T; X<sub>78</sub> is E or D; X<sub>79</sub> is Q or E; X<sub>80</sub> is S, R, T, or N; X<sub>83</sub> is E or D; X<sub>85</sub> is F or L; X<sub>86</sub> is S or Y; X<sub>88</sub> is E or G; X<sub>90</sub> is Y, H, N; X<sub>95</sub> is D, E, or N; X<sub>101</sub> is I, M, or V; X<sub>103</sub> is E or G; X<sub>105</sub> is G or W; X<sub>106</sub> is V or M; X<sub>107</sub> is E, G, or K; X<sub>108</sub> is E or G; X<sub>114</sub> is V, E, or G; X<sub>116</sub> is S or P; X<sub>121</sub> is K or R; X<sub>124</sub> is F or L; X<sub>132</sub> is T, I, or M; X<sub>134</sub> is K or R; and X<sub>140</sub> is A or S; or a fragment of said SEQ ID NO:71. In another aspect, the interferon homologue polypeptide of SEQ ID NO:71, or a fragment thereof, exhibits an antiproliferative activity in a human Daudi cell line-based proliferation assay (at least about  $8.3 \times 10^6$  units/mg) and/or an antiviral activity in a human WISH cell/EMCV-based assay (at least about  $2.1 \times 10^7$  units/mg). Both such assays are discussed in greater detail below. Such polypeptide may comprise an amino acid sequence of the group of from SEQ ID NO:36 to SEQ ID NO:54 or may be encoded by a nucleotide sequence of the group of from SEQ ID NO:1 to SEQ ID NO:19.

Fragments of the interferon homologue polypeptides described herein are also a feature of the invention. An interferon alpha homologue fragment of the invention typically comprises an interferon homologue polypeptide comprising at least about 20, 25, or 30, and typically at least about 40, 50, 60, 70, 80, 90, or 100 contiguous amino acids of  
5 any one of SEQ ID NOS:36-71 or SEQ ID NOS:79-85. In other embodiments, the fragment comprises usually at least about 100, 110, 120, 125, 130, 140, 150, 155, 158, 160, 162, 163, 164, or 165 contiguous amino acids of any one of SEQ ID NOS:36-71 or SEQ ID NOS:79-85. Such polypeptide fragments may have an antiproliferative activity in a human Daudi cell line-based assay and/or an antiviral activity in a human or murine cell  
10 line/EMCV-based assay.

In other embodiments, the invention provides polypeptides having a length of 166 amino acids, and, in some such embodiments, such polypeptides have an antiproliferative activity in a human Daudi cell line-based assay (or other similar assay), including, *e.g.*, at least about  $8.3 \times 10^6$  units/mg, and/or an antiviral activity in a human  
15 WISH cell line/EMCV-based assay (or other similar assay), including, *e.g.*, at least about  $2.1 \times 10^7$  units/mg.

In other embodiments, the invention provides a polypeptide comprising at least 100, 150, 155, or 160 contiguous amino acids of a protein encoded by a coding polynucleotide sequence comprising any of the following: (a) SEQ ID NO:1 to SEQ ID  
20 NO:35 or SEQ ID NO:72 to SEQ ID NO:78; (b) a coding polynucleotide sequence that encodes a first polypeptide selected from any of SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85; and (c) a complementary polynucleotide sequence that hybridizes under at least highly stringent (or ultra-high stringent or ultra-ultra- high stringent conditions) hybridization conditions over substantially the entire length of a  
25 polynucleotide sequence of (a) or (b). Such polypeptides may have an antiproliferative activity in a human Daudi cell line-based assay (or other similar assay), and/or an antiviral activity in a human WISH cell line/EMCV-based assay (or other similar assay). Some such polypeptides of the invention specifically bind a human alpha interferon receptor. The polypeptides and nucleic acids of the subject invention need not be identical, but can  
30 be substantially identical, to the corresponding sequence of the target molecule or related molecule, including the polypeptides of any of SEQ ID NOS:36-71 or fragments thereof (including those having antiviral or antiproliferative activities in the assays described

herein), or the nucleic acids of any of SEQ ID NOS:1-35 or fragments thereof (including those having antiviral or antiproliferative activities in the assays described herein). The polypeptides can be subject to various changes, such as insertions, deletions, and substitutions, either conservative or non-conservative, where such changes might provide  
5 for certain advantages in their use. The polypeptides of the invention can be modified in a number of ways so long as they comprise a sequence substantially identical (as defined below) or having a percent identity to a sequence in the naturally occurring or known interferon polypeptide molecule.

Alignment and comparison of relatively short amino acid sequences (less  
10 than about 30 residues) is typically straightforward. Comparison of longer sequences can require more sophisticated methods to achieve optimal alignment of two sequences. Optimal alignment of sequences for aligning a comparison window can be conducted by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.*  
15 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Nat'l Acad. Sci. (USA)* 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection, and the best alignment (*i.e.*, resulting in the highest percentage of sequence similarity over the  
20 comparison window) generated by the various methods is selected.

The term sequence identity means that two polynucleotide sequences are identical (*i.e.*, on a nucleotide-by-nucleotide basis) over a window of comparison. The term "percentage of sequence identity" or "percent sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining  
25 the number of positions at which the identical residues occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. In one aspect, the present invention provides interferon homologue nucleic acids having at least about 80%, 85%,  
30 90%, 95%, 96%, 97%, 98%, 99%, 99.5% or more percent sequence identity with the nucleic acids of any of SEQ ID NOS:1-35 or SEQ ID NOS:72-78 or fragments thereof.

As applied to polypeptides, the term substantial identity means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights (described in detail below), share at least about 80 percent sequence identity, preferably at least about 90 percent sequence identity, more preferably 5 at least about 95 percent sequence identity or more (e.g., 97, 98, or 99 percent sequence identity). Preferably, residue positions which are not identical differ by conservative amino acid substitutions. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a 10 group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and 15 methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. In one aspect, the present invention provides interferon homologue polypeptides having at least about 80%, 85%, 90%, 95%, 96%, 97%, 98% 99% 99.5% or 20 more percent sequence identity with the polypeptides of any of SEQ ID NOS:36-71 or SEQ ID NOS:79-85 or fragments thereof.

A preferred example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the FASTA algorithm, which is described in Pearson, W.R. & Lipman, D.J., 1988, *Proc. Nat'l Acad. Sci. USA* 85: 2444. See also W. R. Pearson, 1996, *Methods Enzymol.* 266: 227-258. Preferred parameters used in a 25 FASTA alignment of DNA sequences to calculate percent identity are optimized, BL50 Matrix 15: -5, k-tuple= 2; joining penalty= 40, optimization= 28; gap penalty -12, gap length penalty =-2; and width= 16.

Another preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 30 algorithms, which are described in Altschul *et al.*, 1977, *Nuc. Acids Res.* 25: 3389-3402 and Altschul *et al.*, 1990, *J. Mol. Biol.* 215: 403-410, respectively. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity

for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence,

5 which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be

10 increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum

15 achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both

20 strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) *Proc. Nat'l Acad. Sci. U.S.A.* 89: 10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul (1993) *Proc. Nat'l Acad. Sci. U.S.A.* 90: 5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

Another example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the clustering relationships used to create the alignment. PILEUP 5 uses a simplification of the progressive alignment method of Feng & Doolittle (1987) *J. Mol. Evol.* 35: 351-360. The method used is similar to the method described by Higgins & Sharp (1989) *CABIOS* 5: 151-153. The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster 10 of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence 15 comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, e.g., version 7.0 (Devereaux *et al.* (1984) *Nuc. Acids Res.* 12: 20 387-395.

Another preferred example of an algorithm that is suitable for multiple DNA and amino acid sequence alignments is the CLUSTALW program (Thompson, J. D. *et al.* (1994) *Nucl. Acids. Res.* 22: 4673-4680). ClustalW performs multiple pairwise comparisons between groups of sequences and assembles them into a multiple alignment 25 based on homology. Gap open and Gap extension penalties were 10 and 0.05, respectively. For amino acid alignments, the BLOSUM algorithm can be used as a protein weight matrix (Henikoff and Henikoff (1992) *Proc. Nat'l Acad. Sci. U.S.A.* 89: 10915-10919).

Making Polypeptides of the Invention  
30 Recombinant methods for producing and isolating interferon homologue polypeptides of the invention are described above. In addition to recombinant production, the polypeptides may be produced by direct peptide synthesis using solid-phase techniques

(cf. Stewart *et al.* (1969) *Solid-Phase Peptide Synthesis*, W.H. Freeman Co, San Francisco; Merrifield, J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Peptide synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin 5 Elmer, Foster City, Calif.) in accordance with the instructions provided by the manufacturer. For example, subsequences may be chemically synthesized separately and combined using chemical methods to provide full-length interferon homologues. Fragments of the interferon homologue polypeptides of the invention, as discussed in greater detail above, are also a feature of the invention and may be synthesized by using 10 the procedures described above.

Polypeptides of the invention can be produced by introducing into a population of cells a nucleic acid of the invention, wherein the nucleic acid is operatively linked to a regulatory sequence effective to produce the encoded polypeptide, culturing the cells in a culture medium to produce the polypeptide, and optionally isolating the 15 polypeptide from the cells or from the culture medium.

In another aspect, polypeptides of the invention can be produced by introducing into a population of cells a recombinant expression vector comprising at least one nucleic acid of the invention, wherein the at least one nucleic acid is operatively linked to a regulatory sequence effective to produce the encoded polypeptide, culturing the 20 cells in a culture medium under suitable conditions to produce the polypeptide encoded by the expression vector, and optionally isolating the polypeptide from the cells or from the culture medium.

#### Using Polypeptides

##### Antibodies

25 In another aspect of the invention, an interferon homologue polypeptide of the invention is used to produce antibodies which have, *e.g.*, diagnostic, prophylactic and therapeutic uses, *e.g.*, related to the activity, distribution, and expression of interferon homologues.

Antibodies to interferon homologues of the invention may be generated by 30 methods well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments

produced by an Fab expression library. Antibodies, *i.e.*, those which block receptor binding, are especially preferred for therapeutic or prophylactic use.

Interferon homologue polypeptides for antibody induction do not require biological activity; however, the polypeptide or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least 10 amino acids, preferably at least 15 or 20 amino acids. Short stretches of an interferon homologue polypeptide may be fused with another protein, such as keyhole limpet hemocyanin, and antibody produced against the chimeric molecule.

Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art, and many antibodies are available. *See, e.g.*, Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY; Stites *et al.* (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) *Nature* 256:495-497. Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. *See, Huse et al.* (1989) *Science* 246:1275-1281; and Ward *et al.* (1989) *Nature* 341:544-546. Specific monoclonal and polyclonal antibodies and antisera will usually bind with a  $K_D$  of at least about 0.1  $\mu\text{M}$ , preferably at least about 0.01  $\mu\text{M}$  or better, and most typically and preferably, 0.001  $\mu\text{M}$  or better.

Detailed methods for preparation of chimeric (humanized) antibodies can be found in U.S. Patent 5,482,856. Additional details on humanization and other antibody production and engineering techniques can be found in Borrebaeck (ed.) (1995) *Antibody Engineering*, 2<sup>nd</sup> Edition Freeman and Company, NY (Borrebaeck); McCafferty *et al.* (1996) *Antibody Engineering, A Practical Approach*, IRL at Oxford Press, Oxford, England (McCafferty), and Paul (1995) *Antibody Engineering Protocols*, Humana Press, Towata, NJ (Paul).

In one useful embodiment, this invention provides for fully humanized antibodies against the interferon homologues of the invention. Humanized antibodies are especially desirable in applications where the antibodies are used as prophylactics and therapeutics *in vivo* and *ex vivo* in human patients. Human antibodies consist of

characteristically human immunoglobulin sequences. The human antibodies of this invention can be produced in using a wide variety of methods (*see, e.g.*, Larrick *et al.*, U.S. Pat. No. 5,001,065, and Borrebaeck McCafferty and Paul, *supra*, for a review). In one embodiment, the human antibodies of the present invention are produced initially in 5 trioma cells. Genes encoding the antibodies are then cloned and expressed in other cells, such as nonhuman mammalian cells. The general approach for producing human antibodies by trioma technology is described by Ostberg *et al.* (1983), *Hybridoma* 2:361-367, Ostberg, U.S. Pat. No. 4,634,664, and Engelman *et al.*, U.S. Pat. No. 4,634,666. The antibody-producing cell lines obtained by this method are called triomas because they are 10 descended from three cells; two human and one mouse. Triomas have been found to produce antibody more stably than ordinary hybridomas made from human cells.

#### Adjuvants

In one aspect, the interferon homologue polypeptides of the present invention or fragments thereof are useful as adjuvants to stimulate, enhance, potentiate, or 15 augment an immune response related to an antigen when administered together with the antigen or after or before delivery of the antigen. In another aspect, the invention provides methods for administering one or more of the polypeptides invention described herein to a subject.

#### Therapeutic and Prophylactic Agents

20 As described in greater detail below, the interferon homologue polypeptides of the present invention or fragments thereof are useful in the prophylactic and/or therapeutic treatment of a variety of diseases, disorders, or medical conditions.

For example, the invention provides interferon-alpha homologue 25 polypeptides (and interferon-alpha homologue nucleic acids which encode such polypeptides) that have both antiviral and antiproliferative activities in the assays described herein. In one aspect, the invention provides interferon-alpha homologue polypeptides (and interferon-alpha homologue nucleic acids which encode such polypeptides) in which the ratio of antiviral activity to antiproliferative activity is greater than that of other known interferon-alphas such as those listed in GenBank as noted 30 herein. Such polypeptides (and nucleic acids encoding them) are useful in the therapeutic and/or prophylactic treatment of various diseases and disorders, such as, *e.g.*, treatment regimens for hepatitis B, hepatitis C, HIV, and HSV. In such treatment regimens, some

such polypeptides (and nucleic acids encoding them), such as interferon-alpha homologue 2BA8, offer significant advantages over known interferon-alpha compounds, since they likely exhibit lower side effects upon administration than known interferon-alpha compounds, such as interferon-alpha 2a, are of higher potency, and thus may require in 5 lower dosing and cause fewer immunogenicity effects.

### SEQUENCE VARIATIONS

#### Conservatively Modified Variations

Interferon homologue polypeptides of the present invention include one or more conservatively modified variations (or "conservative variations" or conservative 10 substitutions") of the polypeptide sequences disclosed herein as SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID NO:79 to SEQ ID NO:85. Such conservatively modified variations comprise substitutions, additions or deletions which alter, add or delete a single amino acid or a small percentage of amino acids (typically less than about 5%, more typically less than about 4%, 2%, or 1%) in any of SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID 15 NO:79 to SEQ ID NO:85.

For example, a conservatively modified variation (e.g., deletion) of the 166 amino acid polypeptide identified herein as SEQ ID NO:36 has a length of at least about 157 or 158 amino acids, preferably at least about 159 or 160 amino acids, more preferably at least about 162 or 163 amino acids, and still more preferably at least about 164 or 165 20 amino acids, corresponding to a deletion of less than about 5%, 4%, 2% or 1% of the polypeptide sequence, respectively.

Another example of a conservatively modified variation (e.g., a "conservatively substituted variation") of the polypeptide identified herein as SEQ ID NO:36 will contain "conservative substitutions", according to the six substitution groups 25 set forth in Table 2 (*supra*), in up to about 8 residues (i.e., less than about 5%) of the 166 amino acid polypeptide.

The interferon homologue polypeptide sequences of the invention, including conservatively substituted sequences, can be present as part of larger polypeptide sequences such as which occur upon the addition of one or more domains for 30 purification of the protein (e.g., poly His segments, FLAG epitope segments, etc.), e.g., where the additional functional domains have little or no effect on the activity of the

interferon-alpha portion of the protein, or where the additional domains can be removed by post synthesis processing steps such as by treatment with a protease.

In another embodiment, interferon homologue polypeptides of the present invention comprise the following sequence, identified herein as SEQ ID NO:71:

- 5 CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof, where X<sub>11</sub> is N or D; X<sub>12</sub> is R, S, or K; X<sub>15</sub> is L or M; X<sub>16</sub> is I, M, or V; X<sub>19</sub> is A or G; X<sub>22</sub> is G or R; X<sub>24</sub> is I or T; X<sub>26</sub> is P or H; X<sub>34</sub> is H, Y or Q; X<sub>38</sub> is F or L; X<sub>40</sub> is Q or R; X<sub>45</sub> is G or S; X<sub>46</sub> is N or H; X<sub>47</sub> is Q or R; X<sub>50</sub> is K or R; X<sub>51</sub> is A or T; X<sub>55</sub> is S or F; X<sub>56</sub> is V or A; X<sub>57</sub> is L or F; X<sub>60</sub> is M or I; X<sub>61</sub> is I or M; X<sub>67</sub> is L or F; X<sub>72</sub> is D or N; X<sub>75</sub> is A or V; X<sub>76</sub> is A or T; X<sub>78</sub> is E or D; X<sub>79</sub> is Q or E; X<sub>80</sub> is S, R, T, or N; X<sub>83</sub> is E or D; X<sub>85</sub> is F or L; X<sub>86</sub> is S or Y; X<sub>88</sub> is E or G; X<sub>90</sub> is Y, H, N; X<sub>95</sub> is D, E, or N; X<sub>101</sub> is I, M, or V; X<sub>103</sub> is E or G; X<sub>105</sub> is G or W; X<sub>106</sub> is V or M; X<sub>107</sub> is E, G, or K; X<sub>108</sub> is E or G; X<sub>114</sub> is V, E, or G; X<sub>116</sub> is S or P; X<sub>121</sub> is K or R; X<sub>124</sub> is F or L; X<sub>132</sub> is T, I, or M; X<sub>134</sub> is K or R; and X<sub>140</sub> is A or S; or a fragment of said SEQ ID NO:71. As defined above, a conservatively modified variation of the sequence of SEQ ID NO:71 can include up to a total of about 8 amino acid deletions, insertions, or conservative substitutions in the 166 amino acid polypeptide, excluding the positions designated X in SEQ ID NO:71, which correspond to the amino acid explicitly defined.

- As an example, if four conservative substitutions were localized in the subsequence corresponding to amino acids 141-166 of SEQ ID NO:71, examples of conservatively substituted variations of this subsequence,

WEVVR AEIMR SFSFS TNLQK RLRRKE, include:

WEVVR SEIMR SFSYS TNLQR RLRRKD and

WELVR AEIVR SFSFS TNLNK RLRKKE, and the like, where the conservative substitutions are underlined.

- 30 A feature of the invention is an interferon homologue polypeptide comprising at least about 20, usually at least about 25, typically at least about 30, 40, 50, 60, 70, 80, 90, or 100 contiguous amino acids of any one of SEQ ID NOS:36-71 or SEQ

ID NOS:79-85. In other embodiments, the polypeptide typically comprises at least about 100, 110, 120, 125, 130, 140, 150, 155, 158, 160, 163, 164, or 165 contiguous amino acids of any one of SEQ ID NOS:36-70 or SEQ ID NOS:79-85.

- In other embodiments, the interferon homologue polypeptide of the
- 5 invention comprises an amino acid sequence comprising one or more of amino acid residues (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, wherein the numbering of the amino acids corresponds to the numbering of amino acids in the amino acid sequence of SEQ ID NO:36. In a preferred embodiment, the interferon homologue polypeptide comprises an amino acid sequence
- 10 comprising at least 150, 155, or 166 contiguous amino acid residues of any one of SEQ ID NOS:36-70, further comprising Lys160 and Glu166, wherein the numbering of the amino acids corresponds to the numbering of amino acids in the amino acid sequence of SEQ ID NO:36. Some such polypeptides also exhibit an antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in the human Daudi cell line - based assay, or an antiviral activity
- 15 of at about least  $2.1 \times 10^7$  units/milligram (mg) in the human WISH cell/EMCV-based assay.

#### DEFINING POLYPEPTIDES BY IMMUNOREACTIVITY

- Because the polypeptides of the invention provide a variety of new polypeptide sequences as compared to other alpha interferon homologues, the
- 20 polypeptides also provide a new structural features which can be recognized, *e.g.*, in immunological assays. The generation of antisera which specifically binds the polypeptides of the invention, as well as the polypeptides which are bound by such antisera, are features of the invention.

- The invention includes interferon-alpha homologue polypeptides that
- 25 specifically bind to or that are specifically immunoreactive with an antibody or antisera generated against an immunogen comprising an amino acid sequence selected from one or more of SEQ ID NO:36 to SEQ ID NO:70, SEQ ID NO:71, and SEQ ID NO:79 to SEQ ID NO:85. To eliminate cross-reactivity with other interferon-alpha polypeptides, *e.g.*, known interferon-alpha polypeptides, the antibody or antisera (or antiserum) is subtracted
- 30 with available known alpha interferons, such as those polypeptides encoded by nucleic acids represented by GenBank accession numbers J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545

(IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), and M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), or any other known interferon-alpha polypeptides (typically referred to as  
5 the "control alpha interferon polypeptides"). Where the accession number corresponds to a nucleic acid, a polypeptide encoded by the nucleic acid is generated and used for antibody/antisera subtraction purposes. Where the nucleic acid corresponds to a non-coding sequence, *e.g.*, a pseudo gene, an amino acid which corresponds to the reading frame of the nucleic acid is generated (*e.g.*, synthetically), or is minimally modified to  
10 include a start codon for recombinant production.

In one typical format, the immunoassay uses a polyclonal antiserum which was raised against one or more polypeptides comprising one or more of the amino acid sequences corresponding to one or more of: SEQ ID NO:36 to SEQ ID NO:70, SEQ ID NO:71, and SEQ ID NO:79 to SEQ ID NO:85, or a substantial subsequence thereof (*i.e.*,  
15 at least about 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 98% or more of the full length sequence provided). The full set of potential polypeptide immunogens derived from one or more of SEQ ID NO:36 to SEQ ID NO:70, SEQ ID NO:7, and SEQ ID NO:79 to SEQ ID NO:85 are collectively referred to below as "the immunogenic polypeptides." The resulting antisera is optionally selected to have low cross-reactivity  
20 against the control alpha interferon polypeptides and/or other known interferon polypeptides and any such cross-reactivity is removed by immunoabsorption with one or more of the control alpha interferon polypeptides, prior to use of the polyclonal antiserum in the immunoassay.

In order to produce antisera for use in an immunoassay, one or more of the  
25 immunogenic polypeptides is produced and purified as described herein. For example, recombinant protein may be produced in a mammalian cell line. An inbred strain of mice (used in this assay because results are more reproducible due to the virtual genetic identity of the mice) is immunized with the immunogenic polypeptide(s) in combination with a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol  
30 (*see* Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a standard description of antibody generation, immunoassay formats and conditions that can be used to determine specific immunoreactivity).

Alternatively, one or more synthetic or recombinant polypeptides derived from the sequences disclosed herein is conjugated to a carrier protein and used as an immunogen.

Polyclonal sera are collected and titered against the immunogenic polypeptide(s) in an immunoassay, for example, a solid phase immunoassay with one or 5 more of the immunogenic polypeptides immobilized on a solid support. Polyclonal antisera with a titer of  $10^6$  or greater are selected, pooled and subtracted with the control alpha interferon polypeptides to produce subtracted pooled titered polyclonal antisera.

The subtracted pooled titered polyclonal antisera are tested for cross reactivity against the control alpha interferon polypeptides. Preferably at least two of the 10 immunogenic alpha interferon polypeptides are used in this determination, preferably in conjunction with at least two of the control alpha interferon polypeptides, to identify antibodies which are specifically bound by the immunogenic polypeptides(s).

In this comparative assay, discriminatory binding conditions are determined for the subtracted titered polyclonal antisera which result in at least about a 5-10 fold 15 higher signal to noise ratio for binding of the titered polyclonal antisera to the immunogenic alpha interferons as compared to binding to the control alpha interferons . That is, the stringency of the binding reaction is adjusted by the addition of non-specific competitors such as albumin or non-fat dry milk, or by adjusting salt conditions, temperature, or the like. These binding conditions are used in subsequent assays for 20 determining whether a test polypeptide is specifically bound by the pooled subtracted polyclonal antisera. In particular, test polypeptides which show at least a 2-5x higher signal to noise ratio than the control polypeptides under discriminatory binding conditions, and at least about a  $\frac{1}{2}$  signal to noise ratio as compared to the immunogenic 25 polypeptide(s), shares substantial structural similarity or homology with the immunogenic polypeptide as compared to known alpha interferons, and is, therefore a polypeptide of the invention.

In another example, immunoassays in the competitive binding format are used for detection of a test polypeptide. For example, as noted, cross-reacting antibodies are removed from the pooled antisera mixture by immunoabsorption with the control alpha 30 interferon polypeptides. The immunogenic polypeptide(s) are then immobilized to a solid support which is exposed to the subtracted pooled antisera. Test proteins are added to the assay to compete for binding to the pooled subtracted antisera. The ability of the test

protein(s) to compete for binding to the pooled subtracted antisera as compared to the immobilized protein(s) is compared to the ability of the immunogenic polypeptide(s) added to the assay to compete for binding (the immunogenic polypeptides compete effectively with the immobilized immunogenic polypeptides for binding to the pooled antisera). The percent cross-reactivity for the test proteins is calculated, using standard calculations.

In a parallel assay, the ability of the control proteins to compete for binding to the pooled subtracted antisera is determined as compared to the ability of the immunogenic polypeptide(s) to compete for binding to the antisera. Again, the percent cross-reactivity for the control polypeptides is calculated, using standard calculations. Where the percent cross-reactivity is at least 5-10x as high for the test polypeptides, the test polypeptides are said to specifically bind the pooled subtracted antisera.

In general, the immunoabsorbed and pooled antisera can be used in a competitive binding immunoassay as described herein to compare any test polypeptide to the immunogenic polypeptide(s). In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the subtracted antisera to the immobilized protein is determined using standard techniques. If the amount of the test polypeptide required is less than twice the amount of the immunogenic polypeptide that is required, then the test polypeptide is said to specifically bind to an antibody generated to the immunogenic polypeptide, provided the amount is at least about 5-10x as high as for a control polypeptide.

As a final determination of specificity, the pooled antisera is optionally fully immunosorbed with the *immunogenic* polypeptide(s) (rather than the control polypeptides) until little or no binding of the resulting immunogenic polypeptide subtracted pooled antisera to the immunogenic polypeptide(s) used in the immunoabsorption is detectable. This fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If little or no reactivity is observed (*i.e.*, no more than 2x the signal to noise ratio observed for binding of the fully immunosorbed antisera to the immunogenic polypeptide), then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

ANTIPROLIFERATIVE PROPERTIES OF INTERFERON HOMOLOGUES

The effect of interferon homologues on cellular growth was examined in a human Daudi cell line - based assay as described in Example 1. Fig. 2 shows the antiproliferative activity of exemplary interferon homologues of the invention comprising 5 amino acid sequences SEQ ID NO:36 to SEQ ID NO:54, in comparison to control interferons, human IFN-alpha 2a and consensus human IFN-alpha (Con1). The graph shows the number of Units of activity per milligram (mg) of interferon test sample (Y axis) for a set of exemplary interferon alpha homologues, each of which is designated with a name (clone name) on the X axis, compared with that of human IFN-alpha 2a and 10 consensus human IFN-alpha. These results indicate that compositions comprising an interferon-alpha homologue of the present invention can be used in methods to inhibit or reduce proliferation of tumor cells, including, but not limited to: human carcinoma cells, hematopoietic cancer cells, human leukemia cells, human lymphoma cells, and human melanoma cells. Inhibition can be performed *in vitro* (useful, e.g., in a variety of 15 proliferation assays), *ex vivo* or *in vivo* (useful, e.g., as a therapeutic or prophylactic agent).

Interferon-alpha homologues of the present invention show diverse activity patterns against a variety of cancer cell lines (*see, e.g.*, Example 2). An *in vitro* cell line screen (as described in, *e.g.*, Monks, A. *et al.* (1991) J. Nat'l Cancer Inst. 83:757-766) was 20 used to assay interferon-alpha homologues of the invention for selective growth inhibition and/or cell killing of particular cancer cell lines. The human cancer cell lines screened (*see, e.g.*, Example 2, Table 3) include leukemias, melanomas, and cancers of the lung, colon, brain, central nervous system, ovary, breast, prostate, and kidney.

Three activity parameters were determined in the cancer cell line screen: 1) 25 GI50 ("growth inhibition at 50%"), a measure of growth inhibition activity, is the concentration of interferon test sample (IFN alpha homologue or control IFN alpha) at which cell growth is inhibited by 50%, as measured by a 50% reduction in the net protein/polypeptide increase in the interferon test sample as compared to that observed in the control cells (no test sample) at the end of the incubation period; 2) TGI ("total growth 30 inhibition") a measure of cytostatic activity, is the concentration of interferon test sample at which cell growth of a particular cell line is totally inhibited, wherein the amount of cellular protein at the end of the incubation period equals the amount of cellular protein at

the beginning of the incubation period ; and 3) LC50, a measure of cytotoxic activity, is the concentration of interferon test sample at which a 50% reduction in the measured amount of cellular protein at the end of the incubation as compared to that at the beginning of the incubation period is observed, indicating a net loss of cells following interferon test  
5 sample addition. Further details of the assay and data analysis procedures are provided in Example 2.

The activity parameters of exemplary interferon-alpha homologue 3DA11 (SEQ ID NO:40) against a variety of cancer cell lines are shown in Figs. 3A, 3B, and 3C, in comparison with the interferon-alpha Con1 and human interferon-alpha 2a controls.

10 With respect to growth inhibition activity, in particular, homologue 3DA11 and control interferon-alpha Con1 showed significant activity against most of the cell lines tested, with the interferon-alpha Con1 exhibiting generally higher activity, and interferon-alpha 2a generally exhibiting lower overall activity and in only a subset of the cell lines (Fig. 3A).

15 In contrast, in particular, a pronounced difference was observed in the cytotoxic and cytostatic activities of homologue 3DA11 in comparison to both interferon-Con1 and human interferon-alpha 2a controls. In the concentration range tested, homologue 3DA11 showed significant cytostatic activity against a population of cells of eleven of the cell lines, while interferon-Con1 showed activity against only a population  
20 of cells of one of the cell lines, against which homologue 3DA11 was also active (Fig. 3B). IFN-alpha 2a, on the other hand, was not active in this assay against any of the tested cell lines. Homologue 3DA11 thus has a broader cytostatic activity profile than consensus human interferon-alpha (Con1) and human interferon-alpha 2a.

Homologue 3DA11 also showed significant cytotoxic activity in  
25 comparison to the interferon-Con1 and human interferon-alpha 2a controls (Fig. 3C). Surprisingly, homologue 3DA11 displayed cytotoxic activity against a population of cells of 8 of the cell lines, whereas neither the interferon-Con1 nor the interferon-alpha 2a controls exhibited measurable activity against a population of cells of any of the cell lines at the concentration range employed in the assay. Thus, homologue 3DA11 also has a  
30 broader cytotoxic activity profile than interferon-Con1 and human interferon-alpha 2a.

Figs. 4A-4D illustrate the cytostatic activity (as reflected by the TGI value) of exemplary interferon-alpha homologues of the invention. In each figure, the relative

cytostatic activity (expressed as -log TGI) against a population of cells of particular cancer cell line is plotted for various interferon-alpha homologues and for the two control interferons (interferon-Con1 and human interferon-alpha 2a).

Of the exemplary homologues tested, homologues 1D3 (SEQ ID NO:54) and 3DA11 (SEQ ID NO:40), but neither of the control interferons, exhibited significant cytostatic activity against a population of cells of leukemia cell line RPMI-8226 over the concentration range of the assay (Fig. 4A). In this example, the 1D3 and 3DA11 homologues showed at least about 25-fold higher cytostatic activity against a population of the cells (corresponding to a difference in TGI of at least about 1.4 log units) than did either of the controls (interferon-Con1 or interferon-alpha 2a) against a population of cells of the leukemia cell line.

Homologues 1D3, 2G5 (SEQ ID NO:45), 6CG3 (SEQ ID NO:52) and 3DA11, but neither of the control interferons, exhibited significant cytostatic activity against lung cancer cell line NCI-H23 (Fig. 4B). In this example, the 1D3, 2G5, 6CG3, and 3DA11 homologues showed at least about 12-fold higher cytostatic activity a population of cells of a lung cancer cell line (corresponding to a difference in TGI of at least about 1.1 log units) than either interferon-Con1 or interferon-alpha 2a against a population of cells of the lung cancer cell line.

Homologues 1D3, 2G5, and 3DA11, but neither of the control interferons, showed significant cytostatic activity against a population of cells of renal cancer cell line ACHN (Fig. 4C). In this example, the 1D3, 2G5, and 3DA11 homologues showed at least about 35-fold higher cytostatic activity a population of cells of said renal cancer cell line (corresponding to a difference in TGI of at least about 1.55 log units) than either interferon-Con1 or interferon-alpha 2a against a population of cells of renal cancer cell line.

Homologues 1D3, 2G5, 3DA11, 2CA5 (SEQ ID NO:42) and 2DB11 (SEQ ID NO:41), and the interferon-Con1 control, but not the interferon alpha-2a control, exhibited significant cytostatic activity against a population of cells of an ovarian cancer cell line OVCAR-3 (Fig. 4D). In this example, homologue 1D3 showed at least about 2-fold higher cytostatic activity (corresponding to a difference in TGI of at least about 0.3 log units) than interferon-Con1, and the 1D3, 2G5, 3DA11, 2CA5, and 2DB11 homologues showed at least about 40-fold higher cytostatic activity (corresponding to a

difference in TGI of at least about 1.6 log units) than interferon-alpha 2a, against respective populations of cells of the ovarian cancer cell line.

From the exemplary data provided herein, it is apparent that interferon-alpha homologues of the invention showed a variety of cytostatic activity profiles, which  
5 differed significantly from those of the interferon-alpha Con1 and interferon alpha-2a.

The present invention includes an interferon-alpha homologue having increased cytostatic activity relative to human interferon-alpha 2a or to consensus human interferon-alpha, Con1. In various embodiments, the interferon-alpha homologue has at least about 2-fold higher cytostatic activity a population of cells of a cancer cell line (*i.e.*,  
10 has a TGI value at least about 2-fold lower) than does human interferon-alpha 2a, or has at least 2-fold higher cytostatic activity than interferon-Con1, against a population of cells of one or more cancer cell lines selected from the following: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a central nervous system (CNS) cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a  
15 prostate cancer cell line; and a renal cancer cell line.

In other embodiments, the interferon-alpha homologue has at least about 5-fold higher cytostatic activity a population of cells of a cancer cell line (*i.e.*, has a TGI value at least about 5-fold lower) than does human interferon-alpha 2a, or has at least about 5-fold higher cytostatic activity than interferon-Con1, against a population of cells  
20 of one or more cancer cell lines selected from the following: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a central nervous system (CNS) cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a prostate cancer cell line; and a renal cancer cell line. In other embodiments, the interferon-alpha homologue has at least about 10-fold higher cytostatic activity a  
25 population of cells of a cancer cell line (*i.e.*, has a TGI value at least about 10-fold lower) than does human interferon-alpha 2a, or has at least about 10-fold higher cytostatic activity than interferon-Con1, against a population of cells of one or more cancer cell lines selected from the following: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a CNS cancer cell line; an ovarian cancer cell line; a breast  
30 cancer cell line; a prostate cancer cell line; and a renal cancer cell line.

The invention includes an interferon-alpha homologue having increased cytotoxic activity relative to human interferon-alpha 2a or relative to interferon-Con1. In

various embodiments, the interferon-alpha homologue has at least about 2-fold higher cytotoxic activity (*i.e.*, has an LC50 value at least about 2-fold lower), at least 5-fold higher cytotoxic activity, or at least 10-fold higher cytotoxic activity, than human interferon-alpha 2a against a population of cells of one or more cancer cell lines selected from the following: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a CNS cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a prostate cancer cell line; and a renal cancer cell line. In other embodiments, the interferon-alpha homologue has at least about 2-fold higher cytotoxic activity (*i.e.*, has an LC50 value at least about 2-fold lower), at least about 5-fold higher cytotoxic activity, or at least about 10-fold higher cytotoxic activity, than interferon-Con1, against a population of cells of at least one cancer cell line selected from: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a CNS cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a prostate cancer cell line; and a renal cancer cell line.

The invention includes an interferon-alpha homologue having increased growth inhibition activity relative to human interferon-alpha 2a or to interferon-Con1. In various embodiments, the interferon-alpha homologue has at least about 2-fold higher growth inhibition activity (*i.e.*, has a GI50 value at least about 2-fold lower), at least about 5-fold higher growth inhibition activity, or at least about 10-fold higher growth inhibition activity, than human interferon-alpha 2a, against a population of cells of one or more cancer cell lines selected from: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a CNS cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a prostate cancer cell line; and a renal cancer cell line. In other embodiments, the interferon-alpha homologue has at least about 2-fold higher growth inhibition activity (*i.e.*, has a GI50 value at least about 2-fold lower), at least about 5-fold higher growth inhibition activity, or at least about 10-fold higher growth inhibition activity, than interferon-Con1, against at least one cancer cell line selected from the following: a leukemia cell line; a melanoma cell line; a lung cancer cell line; a colon cancer cell line; a CNS cancer cell line; an ovarian cancer cell line; a breast cancer cell line; a prostate cancer cell line; and a renal cancer cell line.

The discovery set forth herein that interferons (such as the interferon-alpha homologues described herein) can be evolved, modified, or recombined to display a

variety of activity profiles provides an opportunity for evolving and creating customized and specific interferon homologues for the treatment of a variety of specific diseases or disease conditions, including, e.g., a variety of cancers or related conditions. For example, an interferon homologue of the invention optimized to have increased potency against a  
5 particular target cancer cell type may also be optimized to have (advantageously) reduced toxicity towards a non-target cell(s), and thus may produce lower side effects in the subject to which the homologue is administered (e.g., patient).

The present invention further provides an opportunity to optimize interferon homologues against tumor cells taken from a subpopulation of subjects (e.g.,  
10 mammals or human patients), or even from an individual subject (e.g., mammal or human patient), providing therapeutic or prophylactic treatment tailored to the individual subject. Optimized interferon homologues of the invention may provide therapeutic or prophylactic benefit against cancers or related conditions or other interferon-treatable disorders or conditions which are otherwise unresponsive to currently-available interferons  
15 or to other treatment regimes.

#### ANTIVIRAL PROPERTIES OF INTERFERON HOMOLOGUES

The antiviral activity of interferon homologues of the present invention was evaluated in a human WISH cell/EMCV assay as described in Example 1. Fig. 2 shows the antiviral activity of exemplary interferon homologues of the invention comprising  
20 amino acid sequences SEQ ID NO:36 to SEQ ID NO:54.

Improved *in vitro* antiviral activity of exemplary IFN-alpha homologues of the invention has been shown to be maintained *in vivo* in a murine model system. Two IFN-alpha homologues of the invention, designated CH2.2 and CH2.3 (SEQ ID NOS:84 and 85, respectively), were previously shown to have about 206,000-fold and 138,000-fold  
25 improved antiviral activity, respectively, compared to human IFN-alpha 2a in a murine cell-based assay, as well as significantly higher activity in the same assay as compared to native murine interferons (Chang *et al.* (1999) *Nature Biotechnol.* 17:793-797). As described in Example 3 below, Balb/c mice challenged with a lethal dose of vesicular stomatitis virus (VSV) were administered varying doses of IFN-alpha homologues,  
30 designated CH2.2 and CH2.3, native murine interferon Mu-IFN alpha 4, and human IFN-alpha 2a. The high *in vitro* activity correlated well with the observed *in vivo* activity (Fig. 5). The CH2.2 and CH2.3 homologues were fully effective in protecting mice from the

lethal viral challenge, while the same dosage of the native murine interferon was partially effective and the human IFN-alpha 2a was completely ineffective. These results indicate that compositions comprising interferon homologues of the present invention can be used in methods to inhibit viral replication in subjects infected with viruses including, but not limited to: human immunodeficiency virus (HIV), hepatitis C virus (HCV), herpes simplex virus (HSV), and hepatitis B virus (HBV). Inhibition can be performed *in vitro* (useful, e.g., in a variety of antiviral assays), *ex vivo* (useful e.g., as a therapeutic or prophylactic agent in *ex vivo* methods discussed herein), or *in vivo* (useful, e.g., as a therapeutic or prophylactic agent in *in vivo* methods discussed herein).

10 **INTERFERON HOMOLOGUES IN THE TREATMENT OF AUTOIMMUNE AND OTHER IMMUNE-RELATED DISORDERS**

Compositions of the present invention can be used to therapeutically or prophylactically treat and thereby alleviate a variety of immune system-related disorders characterized by hyper- or hypo-active immune system function or other features. Such disorders include hyperallergenicity and autoimmune disorders, such as multiple sclerosis, type I (insulin dependent) diabetes mellitus, lupus erythematosus, amyotrophic lateral sclerosis, Crohn's disease, rheumatoid arthritis, stomatitis, asthma, allergies, psoriasis and the like.

**THERAPEUTIC AND PROPHYLACTIC COMPOSITIONS**

20 Therapeutic or prophylactic compositions comprising one or more interferon homologue polypeptides or nucleic acids of the invention are tested in appropriate *in vitro*, *ex vivo*, and *in vivo* animal models of disease, to confirm efficacy, tissue metabolism, and to estimate dosages, according to methods well known in the art. In particular, dosages can be determined by activity comparison of the alpha interferon homologues to existing alpha interferon therapeutics or prophylactics, i.e., in a relevant assay. In one aspect, the invention provides methods comprising administering one or more interferon homologue nucleotides or polypeptides of the invention (or fragments thereof) described above to a mammal, including, e.g., a human, primate, mouse, pig, cow, goat, rabbit, rat, guinea pig, hamster, horse, sheep; or a non-mammalian vertebrate such as 25 a bird (e.g., a chicken or duck) or a fish, or invertebrate, as described in greater detail below. Such compositions typically comprise one or more interferon homologue

nucleotides or polypeptides of the invention (or fragments thereof) and an excipient, including, *e.g.*, a pharmaceutically acceptable excipient.

In one aspect, a composition of the invention is produced by digesting one or more nucleic acids of the invention (or fragments thereof) with a restriction

5 endonuclease, an RNase, or a DNase.

In another aspect of the invention, compositions produced by incubating one or more nucleic acids described above in the presence of deoxyribonucleotide triphosphates and a nucleic acid polymerase, *e.g.*, a thermostable polymerase, are provided.

10 The invention also includes compositions comprising two or more nucleic acids described above. The composition may comprise a library of nucleic acids, where the library contains at least about 5, 10, 20, 50, 100, 150, or 200 or more such nucleic acids.

15 Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells. The interferon-alpha homologues of the invention are administered in any suitable manner, preferably with pharmaceutically acceptable carriers. Suitable methods of administering such interferon homologues in the context of the present invention to a patient are available, and, although more than one route can be used to administer a particular composition, a particular route 20 can often provide a more immediate and more effective reaction than another route.

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions of the present invention.

25 Polypeptide compositions can be administered for any of the prophylactic, therapeutic, and diagnostic methods described herein by a number of routes including, but not limited to oral, intravenous, intraperitoneal, intramuscular, transdermal, subcutaneous, topical, sublingual, vaginal, or rectal means, or by inhalation. Interferon homologue polypeptide compositions can also be administered via liposomes. Such administration 30 routes and appropriate formulations are generally known to those of skill in the art.

The interferon homologue polypeptide or nucleic acid, alone or in combination with other suitable components, can also be made into aerosol formulations

(*i.e.*, they can be “nebulized”) to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

- Formulations suitable for parenteral administration, such as, for example,
- 5 by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening
- 10 agents, stabilizers, and preservatives. The formulations of packaged nucleic acid can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials.

Parenteral administration and intravenous administration are preferred methods of administration. In particular, the routes of administration already in use for existing alpha interferon therapeutics or prophylactics, along with formulations in current

15 use, are preferred routes of administration and formulation for the alpha interferon homologue polypeptide and nucleic acids of the invention.

Cells transduced with the interferon homologue nucleic acids as described above in the context of *ex vivo* or *in vivo* therapy can also be administered intravenously or parenterally as described above. It will be appreciated that the delivery of cells to subjects

20 (*e.g.*, human patients) is routine, *e.g.*, delivery of cells to the blood via intravenous or intraperitoneal administration.

The dose of interferon homologue polypeptide or nucleic acid of the invention administered to a subject (*e.g.*, patient), in the context of the present invention is sufficient to effect a beneficial therapeutic or prophylactic response in the subject (*e.g.*, patient) over time, or to inhibit infection by a pathogen, depending on the application. The dose will be determined by the efficacy of the particular vector, or formulation, and the activity interferon homologue employed and the condition of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany

25

30 the administration of a particular vector, formulation, transduced cell type or the like in a particular patient.

In the therapeutic and prophylactic treatment methods of the invention described herein, an effective amount of an interferon-alpha nucleic acid (e.g., DNA or mRNA) of the invention (e.g., nucleic acid dosage) will generally be in the range of, e.g., from about 0.05 microgram/kilogram (kg) to about 50 mg/kg, usually about 0.005-5 mg/kg. However, as will be understood, the effective amount of the nucleic acid (e.g., nucleic acid dosage) and/or polypeptide (e.g., polypeptide dosage) will vary in a manner apparent to those of ordinary skill in the art according to a number of factors, including the activity or potency of the polypeptide, the activity or potency of any nucleic acid construct (e.g., vector, promoter, expression system) to be administered, the disease or condition (e.g., particular cancer) to be treated, and the subject to which or whom the nucleic acid is delivered.

For delivery of some polypeptides, e.g., by delivering nucleic acids encoding such polypeptides, for example, adequate levels of translation and/or expression are achieved with a nucleic acid dosage of, e.g., about 0.005mg/kg to about 5 mg/kg.

Dosages for other polypeptides (and nucleic acids encoding them) having a known biological activity can be readily determined by those of skill in the art according to the factors noted above. Dosages used for other known interferon-alphas for particular diseases provide guidelines for determining dosage and treatment regimen for a nucleic acid or polypeptide of the invention. An effective amount of an interferon-alpha homologue polypeptide may be in the range of from about 1 microgram to about 1 milligram, and more typically from about 1 microgram to about 100 micrograms.

A composition for use in therapeutic and prophylactic treatment methods of the invention described herein may comprise, e.g., a concentration of an interferon-alpha homologue nucleic acid (e.g., DNA or mRNA) of the invention of from about 0.1 microgram/milliliter (ml) to about 20 mg/ml and a pharmaceutically acceptable carrier (e.g., aqueous carrier).

A composition for use in therapeutic and prophylactic treatment methods of the invention described herein may comprise, e.g., a concentration of an interferon-alpha homologue polypeptide of the invention in an amount as described above and herein and a pharmaceutically acceptable carrier (e.g., aqueous carrier).

In determining the effective amount of the vector, cell type, or formulation to be administered in the treatment or prophylaxis of cancers or viral diseases, the

physician evaluates circulating plasma levels, vector/cell/formulation/ interferon homologue toxicities, progression of the disease, and the production of anti-vector/interferon homologue antibodies.

The dose administered, *e.g.*, to a 70 kilogram patient will be in the range  
5 equivalent to dosages of currently-used interferon-alpha therapeutic or prophylactic proteins, and doses of vectors or cells which produce interferon homologue sequences are calculated to yield an equivalent amount of interferon homologue nucleic acid or expressed protein. The vectors of this invention can supplement treatment of cancers and virally-mediated conditions by any known conventional therapy, including cytotoxic  
10 agents, nucleotide analogues (*e.g.*, when used for treatment of HIV infection), biologic response modifiers, and the like.

For administration, interferon homologues and transduced cells of the present invention can be administered at a rate determined by the LD-50 of the interferon homologue polypeptide or nucleic acid, vector, or transduced cell type, and the side-  
15 effects of the interferon homologue polypeptides or nucleic acids, vector or cell type at various concentrations, as applied to the mass and overall health of the patient.

Administration can be accomplished via single or divided doses.

For introduction of recombinant alpha-interferon nucleic acid transduced cells into a subject (*e.g.*, patient), blood samples are obtained prior to infusion, and saved  
20 for analysis. Between  $1 \times 10^6$  and  $1 \times 10^{12}$  transduced cells are infused intravenously over 60- 200 minutes. Vital signs and oxygen saturation by pulse oximetry are closely monitored. Blood samples are obtained 5 minutes and 1 hour following infusion and saved for subsequent analysis. Leukopheresis, transduction and reinfusion are optionally repeated every 2 to 3 months for a total of 4 to 6 treatments in a one year period. After the  
25 first treatment, infusions can be performed on a outpatient basis at the discretion of the clinician. If the reinfusion is given as an outpatient, the participant is monitored for at least 4, and preferably 8 hours following the therapy. Transduced cells are prepared for reinfusion according to established methods. See Abrahamsen *et al.* (1991) *J. Clin. Apheresis* 6:48-53; Carter *et al.* (1988) *J. Clin. Arpheresis* 4:113-117; Aebersold *et al.*  
30 (1988), *J. Immunol. Methods* 112:1-7; Mual *et al.* (1987) *J. Immunol. Methods* 101:171-181 and Carter *et al.* (1987) *Transfusion* 27:362-365. After a period of about 2-4 weeks in culture, the cells should number between  $1 \times 10^6$  and  $1 \times 10^{12}$ . In this regard, the growth

characteristics of cells vary from patient to patient and from cell type to cell type. About 72 hours prior to reinfusion of the transduced cells, an aliquot is taken for analysis of phenotype, and percentage of cells expressing the therapeutic or prophylactic agent.

If a subject (*e.g.*, patient) undergoing infusion of a vector or transduced cell  
5 or protein formulation develops fevers, chills, or muscle aches, he/she receives the appropriate dose of aspirin, ibuprofen, acetaminophen or other pain/fever controlling drug. Subjects (*e.g.*, patients) who experience reactions to the infusion such as fever, muscle aches, and chills are premedicated 30 minutes prior to the future infusions with either aspirin, acetaminophen, or, *e.g.*, diphenhydramine. Meperidine is used for more severe  
10 chills and muscle aches that do not quickly respond to antipyretics and antihistamines. Cell infusion is slowed or discontinued depending upon the severity of the reaction.

#### THERAPEUTIC AND PROPHYLACTIC TREATMENT METHODS

The present invention also includes methods of therapeutically or prophylactically treating a disease or disorder by administering *in vivo* or *ex vivo* one or  
15 more nucleic acids or polypeptides of the invention described above (or compositions comprising a pharmaceutically acceptable excipient and one or more such nucleic acids or polypeptides) to a subject, including, *e.g.*, a mammal, including, *e.g.*, a human, primate, mouse, pig, cow, goat, rabbit, rat, guinea pig, hamster, horse, sheep; or a non-mammalian vertebrate such as a bird (*e.g.*, a chicken or duck) or a fish, or invertebrate.

In one aspect of the invention, in *ex vivo* methods, one or more cells or a population of cells of interest of the subject (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) are obtained or removed from the subject and contacted with an amount of a polypeptide of  
25 the invention that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition. The contacted cells are then returned or delivered to the subject to the site from which they were obtained or to another site (*e.g.*, including those defined above) of interest in the subject to be treated. If desired, the contacted cells may be grafted onto a tissue, organ, or system site (including all described above) of interest in  
30 the subject using standard and well-known grafting techniques or, *e.g.*, delivered to the blood or lymph system using standard delivery or transfusion techniques.

The invention also provides *in vivo* methods in which one or more cells or a population of cells of interest of the subject are contacted directly or indirectly with an amount of a polypeptide of the invention effective in prophylactically or therapeutically treating the disease, disorder, or other condition. In direct contact/administration formats, 5 the polypeptide is typically administered or transferred directly to the cells to be treated or to the tissue site of interest (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) by any of a variety of formats, including topical administration, injection (*e.g.*, by using a needle or syringe), or 10 vaccine or gene gun delivery, pushing into a tissue, organ, or skin site. The polypeptide can be delivered, for example, intramuscularly, intradermally, subdermally, subcutaneously, orally, intraperitoneally, intrathecally, intravenously, or placed within a cavity of the body (including, *e.g.*, during surgery), or by inhalation or vaginal or rectal administration.

15 In *in vivo* indirect contact/administration formats, the polypeptide is typically administered or transferred indirectly to the cells to be treated or to the tissue site of interest, including those described above (such as, *e.g.*, skin cells, organ systems, lymphatic system, or blood cell system, *etc.*), by contacting or administering the polypeptide of the invention directly to one or more cells or population of cells from 20 which treatment can be facilitated. For example, tumor cells within the body of the subject can be treated by contacting cells of the blood or lymphatic system, skin, or an organ with a sufficient amount of the polypeptide such that delivery of the polypeptide to the site of interest (*e.g.*, tissue, organ, or cells of interest or blood or lymphatic system within the body) occurs and effective prophylactic or therapeutic treatment results. Such 25 contact, administration, or transfer is typically made by using one or more of the routes or modes of administration described above.

In another aspect, the invention provides *ex vivo* methods in which one or more cells of interest or a population of cells of interest of the subject (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, 30 mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) are obtained or removed from the subject and transformed by contacting said one or more cells or population of cells with a polynucleotide construct

comprising a target nucleic acid sequence of the invention that encodes a biologically active polypeptide of interest (*e.g.*, a polypeptide of the invention) that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition. The one or more cells or population of cells is contacted with a sufficient amount of the  
5 polynucleotide construct and a promoter controlling expression of said nucleic acid sequence such that uptake of the polynucleotide construct (and promoter) into the cell(s) occurs and sufficient expression of the target nucleic acid sequence of the invention results to produce an amount of the biologically active polypeptide effective to prophylactically or therapeutically treat the disease, disorder, or condition. The polynucleotide construct  
10 may include a promoter sequence (*e.g.*, CMV promoter sequence) that controls expression of the nucleic acid sequence of the invention and/or, if desired, one or more additional nucleotide sequences encoding at least one or more of another polypeptide of the invention, a cytokine, adjuvant, or co-stimulatory molecule, or other polypeptide of interest.  
15

Following transfection, the transformed cells are returned, delivered, or transferred to the subject to the tissue site or system from which they were obtained or to another site (*e.g.*, tumor cells, tumor tissue sample, organ cells, blood cells, cells of the skin, lung, heart, muscle, brain, mucosae, liver, intestine, spleen, stomach, lymphatic system, cervix, vagina, prostate, mouth, tongue, *etc.*) to be treated in the subject. If  
20 desired, the cells may be grafted onto a tissue, skin, organ, or body system of interest in the subject using standard and well-known grafting techniques or delivered to the blood or lymphatic system using standard delivery or transfusion techniques. Such delivery, administration, or transfer of transformed cells is typically made by using one or more of the routes or modes of administration described above. Expression of the target nucleic  
25 acid occurs naturally or can be induced (as described in greater detail below) and an amount of the encoded polypeptide is expressed sufficient and effective to treat the disease or condition at the site or tissue system.

In another aspect, the invention provides *in vivo* methods in which one or more cells of interest or a population of cells of the subject (*e.g.*, including those cells and  
30 cells systems and subjects described above) are transformed in the body of the subject by contacting the cell(s) or population of cells with (or administering or transferring to the cell(s) or population of cells using one or more of the routes or modes of administration

described above) a polynucleotide construct comprising a nucleic acid sequence of the invention that encodes a biologically active polypeptide of interest (e.g., a polypeptide of the invention) that is effective in prophylactically or therapeutically treating the disease, disorder, or other condition.

- 5       The polynucleotide construct can be directly administered or transferred to cell(s) suffering from the disease or disorder (e.g., by direct contact using one or more of the routes or modes of administration described above). Alternatively, the polynucleotide construct can be indirectly administered or transferred to cell(s) suffering from the disease or disorder by first directly contacting non-diseased cell(s) or other diseased cells using  
10      one or more of the routes or modes of administration described above with a sufficient amount of the polynucleotide construct comprising the nucleic acid sequence encoding the biologically active polypeptide, and a promoter controlling expression of the nucleic acid sequence, such that uptake of the polynucleotide construct (and promoter) into the cell(s) occurs and sufficient expression of the nucleic acid sequence of the invention results to  
15      produce an amount of the biologically active polypeptide effective to prophylactically or therapeutically treat the disease or disorder, and whereby the polynucleotide construct or the resulting expressed polypeptide is transferred naturally or automatically from the initial delivery site, system, tissue or organ of the subject's body to the diseased site, tissue, organ or system of the subject's body (e.g., via the blood or lymphatic system).  
20      Expression of the target nucleic acid occurs naturally or can be induced (as described in greater detail below) such that an amount of the encoded polypeptide is expressed sufficient and effective to treat the disease or condition at the site or tissue system. The polynucleotide construct may include a promoter sequence (e.g., CMV promoter sequence) that controls expression of the nucleic acid sequence and/or, if desired, one or  
25      more additional nucleotide sequences encoding at least one or more of another polypeptide of the invention, a cytokine, adjuvant, or co-stimulatory molecule, or other polypeptide of interest.

- In each of the *in vivo* and *ex vivo* treatment methods as described above, a composition comprising an excipient and the polypeptide or nucleic acid of the invention  
30      can be administered or delivered. In one aspect, a composition comprising a pharmaceutically acceptable excipient and a polypeptide or nucleic acid of the invention is

administered or delivered to the subject as described above in an amount effective to treat the disease or disorder.

In another aspect, in each *in vivo* and *ex vivo* treatment method described above, the amount of polynucleotide administered to the cell(s) or subject can be an 5 amount sufficient that uptake of said polynucleotide into one or more cells of the subject occurs and sufficient expression of said nucleic acid sequence results to produce an amount of a biologically active polypeptide effective to enhance an immune response in the subject, including an immune response induced by an immunogen (*e.g.*, antigen). In another aspect, for each such method, the amount of polypeptide administered to cell(s) or 10 subject can be an amount sufficient to enhance an immune response in the subject, including that induced by an immunogen (*e.g.*, antigen).

In yet another aspect, in an *in vivo* or *ex vivo* treatment method in which a polynucleotide construct (or composition comprising a polynucleotide construct) is used to deliver a physiologically active polypeptide to a subject, the expression of the 15 polynucleotide construct can be induced by using an inducible on- and off-gene expression system. Examples of such on- and off-gene expression systems include the Tet-On™ Gene Expression System and Tet-Off™ Gene Expression System (*see, e.g.*, Clontech Catalog 2000, pg. 110-111 for a detailed description of each such system), respectively. Other controllable or inducible on- and off-gene expression systems are known to those of 20 ordinary skill in the art. With such system, expression of the target nucleic acid of the polynucleotide construct can be regulated in a precise, reversible, and quantitative manner. Gene expression of the target nucleic acid can be induced, for example, after the stable transfected cells containing the polynucleotide construct comprising the target nucleic acid are delivered or transferred to or made to contact the tissue site, organ or system of 25 interest. Such systems are of particular benefit in treatment methods and formats in which it is advantageous to delay or precisely control expression of the target nucleic acid (*e.g.*, to allow time for completion of surgery and/or healing following surgery; to allow time for the polynucleotide construct comprising the target nucleic acid to reach the site, cells, system, or tissue to be treated; to allow time for the graft containing cells transformed with 30 the construct to become incorporated into the tissue or organ onto or into which it has been spliced or attached, *etc.*)

INTEGRATED SYSTEMS

The present invention provides computers, computer readable media and integrated systems comprising character strings corresponding to the sequence information herein for the polypeptides and nucleic acids herein, including, e.g., those sequences listed 5 herein and the various silent substitutions and conservative substitutions thereof.

Various methods and genetic algorithms (GOs) known in the art can be used to detect homology or similarity between different character strings, or can be used to perform other desirable functions such as to control output files, provide the basis for making presentations of information including the sequences and the like. Examples 10 include BLAST, discussed *supra*.

Thus, different types of homology and similarity of various stringency and length can be detected and recognized in the integrated systems herein. For example, many homology determination methods have been designed for comparative analysis of sequences of biopolymers, for spell-checking in word processing, and for data retrieval 15 from various databases. With an understanding of double-helix pair-wise complement interactions among 4 principal nucleobases in natural polynucleotides, models that simulate annealing of complementary homologous polynucleotide strings can also be used as a foundation of sequence alignment or other operations typically performed on the character strings corresponding to the sequences herein (e.g., word-processing 20 manipulations, construction of figures comprising sequence or subsequence character strings, output tables, etc.). An example of a software package with GOs for calculating sequence similarity or homology is BLAST, which can be adapted to the present invention by inputting character strings corresponding to the sequences herein.

Similarly, standard desktop applications such as word processing software 25 (e.g., Microsoft Word™ or Corel WordPerfect™) and database software (e.g., spreadsheet software such as Microsoft Excel™, Corel Quattro Pro™, or database programs such as Microsoft Access™ or Paradox™) can be adapted to the present invention by inputting a character string corresponding to the interferon alpha homologues of the invention (either nucleic acids or proteins, or both). For example, the integrated systems can include the 30 foregoing software having the appropriate character string information, e.g., used in conjunction with a user interface (e.g., a GUI in a standard operating system such as a Windows, Macintosh or LINUX system) to manipulate strings of characters. As noted,

specialized alignment programs such as BLAST can also be incorporated into the systems of the invention for alignment of nucleic acids or proteins (or corresponding character strings).

- Integrated systems for analysis in the present invention typically include a
- 5 digital computer with GO software for aligning sequences, as well as data sets entered into the software system comprising any of the sequences herein. The computer can be, *e.g.*, a PC (Intel x86 or Pentium chip- compatible DOS™, OS2™ WINDOWS™ NTT™, WINDOWS95™, WINDOWS98™ LINUX based machine, a MACINTOSH™, Power PC, or a UNIX based (*e.g.*, SUN™ work station) machine) or other commercially
- 10 common computer which is known to one of skill. Software for aligning or otherwise manipulating sequences is available, or can easily be constructed by one of skill using a standard programming language such as Visualbasic, Fortran, Basic, Java, or the like.

- Any controller or computer optionally includes a monitor which is often a cathode ray tube ("CRT") display, a flat panel display (*e.g.*, active matrix liquid crystal display, liquid crystal display), or others. Computer circuitry is often placed in a box which includes numerous integrated circuit chips, such as a microprocessor, memory, interface circuits, and others. The box also optionally includes a hard disk drive, a floppy disk drive, a high capacity removable drive such as a writeable CD-ROM, and other common peripheral elements. Inputting devices such as a keyboard or mouse optionally provide for input from a user and for user selection of sequences to be compared or otherwise manipulated in the relevant computer system.

- The computer typically includes appropriate software for receiving user instructions, either in the form of user input into a set parameter fields, *e.g.*, in a GUI, or in the form of preprogrammed instructions, *e.g.*, preprogrammed for a variety of different specific operations. The software then converts these instructions to appropriate language for instructing the operation of the fluid direction and transport controller to carry out the desired operation.

- The software can also include output elements for controlling nucleic acid synthesis (*e.g.*, based upon a sequence or an alignment of a sequences herein) or other operations which occur downstream from an alignment or other operation performed using a character string corresponding to a sequence herein.

In one embodiment, the invention provides an integrated system comprising a computer or computer readable medium comprising a database having one or more sequence records. Each of the sequence records comprises one or more character strings corresponding to a nucleic acid or polypeptide or protein sequence selected from SEQ ID NO:1 to SEQ ID NO:85. The integrated system further comprises a user input interface allowing a user to selectively view the one or more sequence records. In one such integrated system, the computer or computer readable medium comprises an alignment instruction set that aligns the character strings with one or more additional character strings corresponding to a nucleic acid or polypeptide or protein sequence.

One such integrated system includes an instruction set that comprises at least one of the following: a local homology comparison determination, a homology alignment determination, a search for similarity determination, and a BLAST determination. In some embodiments, the system further comprises a readable output element that displays an alignment produced by the alignment instruction set. In another embodiment, the computer or computer readable medium further comprises an instruction set that translates at least one nucleic acid sequence which comprises a sequence selected from SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78 into an amino acid sequence. The instruction set may select the nucleic acid by applying a codon usage instruction set or an instruction set which determines sequence identity to a test nucleic acid sequence.

Methods of using a computer system to present information pertaining to at least one of a plurality of sequence records stored in a database are also provided. Each of the sequence records comprises at least one character string corresponding to SEQ ID NO:1 to SEQ ID NO:85. The method comprises determining at least one character string corresponding to one or more of SEQ ID NO:1 to SEQ ID NO:85 or a subsequence thereof; determining which of the at least one character string of the list are selected by a user; and displaying each of the selected character strings, or aligning each of the selected character strings with an additional character string. The method may further comprise displaying an alignment of each of the selected character strings with an additional character string and/or displaying the list.

KITS

In an additional aspect, the present invention provides kits embodying the methods, composition, systems and apparatus herein. Kits of the invention optionally comprise one or more of the following: (1) an apparatus, system, system component or apparatus component as described herein; (2) instructions for practicing the methods described herein, and/or for operating the apparatus or apparatus components herein and/or for using the compositions herein; (3) one or more alpha interferon homologue compositions (such as *e.g.*, compositions comprising at least one interferon alpha homologue nucleic acid or polypeptide or fragment thereof, cell, vector, *etc.*, of the invention) or components (interferon alpha homologue nucleic acid or polypeptide or fragment thereof, cell, vector, *etc.*, of the invention); (4) a container for holding one or more aspects of the invention, including such components or compositions, and (5) packaging materials.

In a further aspect, the present invention provides for the use of any apparatus, apparatus component, composition or kit herein, for the practice of any method or assay herein, and/or for the use of any apparatus or kit to practice any assay or method herein.

**EXAMPLES****EXAMPLE I: PREPARATION AND SCREENING OF SHUFFLED INTERFERON-ALPHA LIBRARIES**

Fragments (25-60 base pairs (bp) in length) of about 20 human interferon-alpha subspecies genes were prepared by PCR amplification and DNase treatment, and recombined essentially as described in Crameri A. *et al.* (1998; Nature 15:288-291), to produce shuffled interferon-alpha mature coding sequences. Expression libraries were prepared by subcloning shuffled interferon-alpha mature coding sequences into an *E. coli* secretion vector. Shuffled interferon polypeptides were expressed as mature proteins fused at the C-termini to an E tag (Amersham-Pharmacia) to facilitate quantitation and purification from the periplasmic space. *E. coli* transformants were picked using a robotic colony picker (Q-Bot, Genetix Pharmaceuticals) into microtiter plates, and periplasmic extracts were prepared.

Periplasmic extracts were assayed for antiproliferative activity on a human Daudi cell line as described by Scarozza, A.M. *et al.* (1992) *J. Interferon Res.* 12:35-42.

Clones exhibiting antiproliferative activity in the Daudi assay were re-screened and expression levels determined by Western blot using an anti-E tag antibody (Amersham-Pharmacia). Clones exhibiting highest activity normalized to expression levels were selected for sequencing and were also utilized as substrates for additional 5 rounds of shuffling and screening as described above.

Clones from the first and second rounds of shuffling having relatively high antiproliferative activity by the Daudi assay were subcloned into a CHO expression vector (pDEI-1011) in which the E-tag/6-His tag (Amersham-Pharmacia) is fused to the C-terminus of the shuffled interferons. Clones were transfected into CHO cells and stable 10 cell lines were selected with 1 mg/ml G418. CHO-expressed mature interferons were purified on anti-E tag Sepharose column (Amersham-Pharmacia) and quantitated by a Bradford assay (Biorad). CHO-purified shuffled interferons were assayed for antiproliferative activity by the Daudi assay and for antiviral activity using a human WISH cell/EMCV assay as described below.

15 Human WISH cell / EMCV antiviral assay  
WISH cells were seeded to a density of  $6 \times 10^4$  cells/well in 96-well plates in 100  $\mu$ l RPMI medium (Gibco-BRL) supplemented with 10% fetal calf serum, penicillin (100  $\mu$ g/ml), and streptomycin (100  $\mu$ g/ml), and incubated for 24 hours at 37°C. Samples of interferon-alpha polypeptides in medium (100  $\mu$ l total volume) were added to wells and 20 incubated for 3 hours at 37°C under a 5% CO<sub>2</sub> atmosphere. Dilutions of EMCV (encephalomyocarditis virus) were added to wells in 50  $\mu$ l volumes, and incubated for 24 hours as above. Medium was carefully removed and wells were rinsed 2x with warm phosphate-buffered saline (PBS). Neutral red (100  $\mu$ l/well of 1:50 dilution in medium) was added to the wells and incubated for 2 hours as above. Glutaraldehyde (50  $\mu$ l/well of 25 0.5% in PBS) was added and incubated for 30 minutes as above. Wells were washed 2x in PBS, and 100  $\mu$ l/well of a solution of 50% methanol, 1% acetic acid was added. Absorbance at 540 nanometers (nm) was measured using a microplate reader.

Fig. 2 shows the antiproliferative activity and the antiviral activity of exemplary 30 interferon homologues of the invention, in comparison with interferon alpha-2a and interferon-alpha Con1. The graph shows the number of Units activity per milligram of homologue (Y axis) for a set of exemplary interferon alpha homologues, each of which is designated with a "name" on the X axis.

EXAMPLE 2: IN VITRO CANCER CELL LINE SCREEN

An *in vitro* cell line screen (as described in, e.g., Monks, A. *et al.* (1991) *J. Nat'l Cancer Inst.* 83:757-766 (hereinafter "Monks") and <http://dtp.nci.gov./branches/btb/ivclsp.html>, each of which is incorporated herein by reference in its entirety for all purposes) was used to assay interferon-alpha homologues of the invention for selective growth inhibition and/or cell killing of particular cancer cell lines. The 60 human cancer cell lines used (Table 3) include leukemias, melanomas, and cancers of the lung, colon, brain, ovary, breast, prostate, central nervous system, renal system, and kidney. Human tumor cell lines were grown according to procedures outlined in Monks") and <http://dtp.nci.gov./branches/btb/ivclsp.html>.

Table 3  
Human cancer cell lines screened

Cancer type	Cell lines
Leukemia	CCRF-CEM, HL-60 (TB), K-562, MOLT-4, RPMI-8226, SR
Colon cancer	COLO 205, HCC-2998, HCT-15, HCT-116, HT29, KM12, SW-620
CNS cancer	SF-268, SF-295, SF-539, SNB-19, SNB-75, U251
Lung cancer	A549/ATCC, EKVVX, HOP-62, HOP-92, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522
Breast cancer	MCF-7, NCI/ADR HS578T, MDA-MB-231/ATCC, MDA-MB-435, MDA-N, BT-549, T-47D
Melanoma	LOX IMVI, M14, MALME-3M, SK-MEL-2, SK-MEL-5, SK-MEL-28, UACC-62, UACC-257
Ovarian cancer	IGROV1, OVCAR-3, OVCAR-4, OVCAR-5, OVCAR-8, SK-OV-3
Prostate cancer	DU-145, PC-3
Renal cancer	786-0, A498, ACHN, CAKI-1, RXF 393, SN12C, TK-10, UO-31

15            Briefly, cells were inoculated into 96 well microtiter plates at densities ranging from about 5,000 to about 40,000 cells/well, depending on the growth properties of the particular cell line. After inoculation, the microtiter plates were incubated for 24 hours (h) at 37 degrees C prior to addition of test samples (e.g., interferon homologues of the invention or control interferons). After 24 h, two plates of each cell line were fixed *in situ* with trichloroacetic acid (TCA), to provide a measurement of the cell population for each cell line at the time of test sample addition ( $T_0$ ). To the remaining plates, interferon samples (affinity-purified from CHO cell supernatants) were added in five 10-fold serial dilutions ranging from  $10^{-0.8}$  to  $10^{-4.8} \mu\text{g/ml}$ .

Following sample addition, the plates were incubated for an additional 6 days. The assay was terminated by addition of TCA.

Cell population was determined by measuring cellular protein in a quantitative protein dye-binding assay. Sulforhodamine B solution (100  $\mu$ l) at 0.4 % (w/v) in 1% acetic acid was added to each well, followed by incubation for 10 minutes at room temperature. Unbound dye was removed by washing five times with 1% acetic acid and the plates air-dried. Protein-bound dye was solubilized with 10 milliMolar (mM) Tris, and the absorbance read at 515 nanometer (nm) on an automated plate reader .

Seven absorbance measurements were taken for each dose-response assay, 10 corresponding to: the amount of cellular protein prior to sample addition (time zero;  $T_0$ ), the amount of cellular protein at the end of the incubation period in the absence of test sample (control growth, C), and five measurements corresponding to the amount of cellular protein at the end of the incubation period in the presence of each of the five concentrations of interferon test sample (test growth in presence of interferon test sample 15 at the five concentration levels,  $T_i$ ). These measurements were used to calculate the following three parameters for each test sample:

GI50, or "growth inhibition of 50%," is the concentration of interferon test sample at which cell growth is inhibited by 50%, as measured by a 50% reduction in the net protein/polypeptide increase in the interferon test sample as compared to that observed 20 in the control cells (no test sample) at the end of the incubation period. GI50 is calculated as the concentration of test sample where  $[(T_i-T_0)/(C-T_0)] \times 100 = 50$ . See Fig. 3A.

TGI, or "total growth inhibition," is the concentration of interferon test sample at which cell growth is totally inhibited, wherein the amount of cellular protein at the end of the incubation period equals the amount of cellular protein at the beginning of 25 the incubation period. The concentration of interferon test sample that produces total growth inhibition (TGI) is calculated as the concentration of test sample where  $T_i = T_0$ .

LC50 is the concentration of interferon test sample at which a 50% reduction in the measured amount of cellular protein at the end of the incubation as compared to that at the beginning of the incubation period is observed, indicating a net 30 loss of cells following interferon test sample addition. LC50 is calculated as the concentration of test sample where  $[(T_i-T_0)/T_0] \times 100 = -50$ .

If, for a particular test sample, an effect was not achieved or was exceeded at the concentration range tested, the value for that parameter was expressed as greater or less than the maximum or minimum concentration tested.

5    **EXAMPLE 3: IN VITRO ACTIVITY OF IFN-ALPHA HOMOLOGUES CORRELATES WITH IN VIVO EFFICACY.**

Fragments of human interferon-alpha genes were shuffled and screened for activity in a murine cell-based antiviral assay as described by Chang *et al.* (1999) *Nature Biotechnol.* 17:793-797. Interferon-alpha homologues that exhibited over 10<sup>5</sup>-fold higher 10 antiviral activity than human interferon-alpha 2a against mouse cells were isolated. The antiviral activities of a number of the interferon-alpha homologues even significantly exceeded the antiviral activity of native mouse interferons, including Mu-IFN-alpha 4 (Chang *et al.*, *supra*). Recursive sequence recombination (*e.g.*, DNA shuffling) of human interferon-alpha gene fragments to produce novel interferon alpha homologues and 15 subsequent screening of such homologues against murine interferon receptors resulted in the identification and isolation of interferon-alpha homologues with activity optimized for the distantly related murine species.

A dose-response study in mice was performed to determine if the high antiviral activity observed *in vitro* is sustained *in vivo*. Two of the mouse-optimized 20 interferon-alpha homologues, designated herein as CH2.2 and CH2.3 (SEQ ID NOS:84 and 85, respectively), were used in this study. CH2.2 and CH2.3 were shown to have about 138,000-fold and about 206,00-fold higher activity, respectively, than human interferon-alpha 2a, and about 2.5-fold and about 1.6-fold higher activity than native mouse interferon-alpha 4, in the *in vitro* mouse cell antiviral assay (Chang *et al.*, *supra*).

25    Groups of Balb/c mice received subcutaneous doses of either phosphate buffered saline (PBS), interferon-alpha homologue CH2.2, interferon-alpha homologue CH2.3, murine IFN-alpha 4, or human interferon-alpha 2a, in daily subcutaneous doses of 2, 10, or 50 µg (total volume of 50 µl) for four consecutive days. On day 2, the mice were exposed to a lethal intranasal dose (ten times the LC50) of vesicular stomatitis virus 30 (VSV). Data is expressed as the number of mice which survive to day 21.

Fig. 5 shows that both of the mouse-optimized interferon-alpha homologues, CH2.2 and CH2.3, were as effective or more effective than native murine

interferon Mu-IFN alpha 4 in protecting mice from VSV. At the concentrations tested, human IFN-alpha 2a was nearly completely ineffective in protecting mice from the virus. Thus, the *in vivo* efficacy of the interferon-alpha homologues of the invention correlates remarkably well with the antiviral activities observed in the *in vitro* assays.

- 5 While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques, methods, compositions, apparatus and systems described above may be used in various
- 10 combinations. *All publications, patents, patent applications, or other documents cited in this application are incorporated herein by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other document were individually indicated to be incorporated by reference for all purposes.*

#### SEQUENCES

15

SEQ ID	Clone ID	Sequence
SEQ ID NO:1	2DH12	TGTGATCTGCCTCAGACCCACAGCCTGGCAACAGGAGGGCTTGATG CTCCTGGCACAAATGGGACGAATCTCTCCTTCTCCTGCCTGAAGGAC AGACAAGACTTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCATCTGCTGCTTGGAACAGACC CTCCTAGAAAAAATTTCACACTGAACTCTACCAGCAGCTGAATGACTG GAAGCCTGCGTGTACAGGAGGTAGGGGTGAAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGGAAGTACTTCAAAGAATCACT CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:2	2CA3	TGTGATCTGCCTCAGACCCACAGCCTGGTACAGGAGGGCATGATA CTCCTGGCACAAATGGGACGAATCTCTCCTTCTCCTGCCTGAAGGAC AGATATGATTTCGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCATCTGCTGCTTGGAACAGAGC CTCCTAGAAAAAATTTCACACTGAACTTTACCAGCAGCTGAATGAACTG GAAGCATGTGTGTACAGGAGGTGGGGTGGGAGAGACTCCCCTGATG AATGGGACTCCATCCTGGCTGTGAGAAAGTACTTCAAAGAATCACT CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:3	4AB9	TGTGATCTGCCTCAGACCCACAGCCTGGCAACAGGAGGGCTTGATA CTCCTGGCACAAATGGGACGAATCTCTCCTTCTCCTGCCTGAAGGAC AGACATGACTTTGGATTCCCCCGGGAGGAGTTGATGGCAACCAGTTC

		CTCCTGGCACAAATGGGACGAATCTCCTTCTGCCTGAAGGAC AGACATGACTTGATTCCCCGGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAAAGAACTCATCTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATGAAC GAAGCATGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAAC CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:4	2DA4	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGATG CTCCTGGCACAAATGGGAGAAATCTCCTTCTGCCTGAAGGAC AGACAAGACTTGATTCCCCCAGGAGGAGTTGATAGCAACCAGTTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATGAC GAAGCCTGCGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAAC CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:5	3DA11	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGGTA CTCCTGGCACAAATGGGAGAAATCTCCTTCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATGAC GAAGCCTGCGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAAC CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:6	2DB11	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGATG CTCCTGGCACAAATGGGAGAAATCTCCTTCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATGAC GAAGCCTGCGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAAC CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:7	2CA5	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGATA CTCCTGGCACAAATGGGACGAATCTCCTTCTGCCTGAAGGAC AGACAAGACTTGATTCCCCCAGGAGGAGTTGATGGCAACCAGGTT CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGAACTCATCTGCTGCTGGGAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATGAC GAAGCCTGCGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAAC CTTTATCTAATAGAGAGGAAATACAGCCCTTGTGCATGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA

		GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACCCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATCTTCAAAGAACATCACT CTTTATCTAATAGAGAGGAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:8	2G6	TGTGATCTGCCTCAGACCCACAGCCTTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGTCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTAACCAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCCTGATG AATGTGGACCCATCCTGGCTGTGAAGAAATCTTCAAAGAACATCACT CTCTATCTGACAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:9	3AH7	TGTGATCTGCCTCAGACCCACAGCCTTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGCGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATAGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGTCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTCACCAGCAACTGAATGAACCTG GAAGCATGTGTAGTACAGGAGGTTGGGTGGAAGAGACTCCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACCTTCAAAGAACATCACT CTTATCTGACAGAGAAGAAATAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:10	2G5	TGTGATCTGCCTCAGACCCACAGCCTTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACAAGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGTCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTACCAAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACCCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGGAAGTACTTCAAAGAACATCACT CTTATCTAATAGAGAGGAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:11	2BA8	TGTGATCTGCCTCAGACCCACAGCCTTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGGGACGAATCTCTCCTTCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGTCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCACTGCTGCTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTACCAAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACCCCCCTAATG AATGTGGACTCCATCCTGGCTGTGAGGAAGTACTTCAAAGAACATCACT CTTATCTAATAGAGAGGAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA

SEQ ID NO:12	1F3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGGACAAATGGGAAGAATCTCTCATTTCTCCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAACCTCTTCAGCACAAAGGACTCATCTGCTGCTGGATGAGAGG CTTCTAGACAAACTCTATACTGAACCTTACCAAGCAGCTGAATGACCTG GAAGCCTGTGTGATGCAGGAGGTGTTGGGGACTCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGAAAATACTTCCAAGAATCACT CTCTATCTGACAGAGAAGAATAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:13	4BE10	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGCACAGATGGGACGAATCTCCTTCTCCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATAATGCAGCAGACC TTCATCTCTTCAGCACAAAGAAGTCACTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAGACTGAATGACCTG GAAGCATGTGTGATACAGGGGGTTGGGGTGGAAAGAGACTCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGAAAATACTTCCAAGAATCACT CTTATCTGACAGAGAAGAATAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:14	2DD9	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATG CTCCTGGCACAAATGGGAAGAATCTCCCTTCTCCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCATCTCTTCAGCACAAAGGATTCACTGCTGCTGGAACAGAGC CTCCTAGAAAAATTTCACACTGGACTCTACCAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGGGGTTGGGGTGGAAAGAGACCCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGAAAATACTTCCAAGAATCACT CTTATCTGACAGAGAAGAATAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:15	3CA1	TGTGATCTGCCTCAGACCCACAGCCTGGCAACAGGAGGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCTCTTCTCCTGCCTGAAGGAC AGACATGACTTGGATTACCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCATCTCTTCAGCACAAAGAAGTCACTGCTGCTGGATGAGACC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAGACTGAATAACCTG GAAGCATGTGTGATACAGGGGGTTGGGGTGGAAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGAAAATACTTCCAAGAATCACT CTTATCTGACAGAGAAGAATAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:16	2F8	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGCACAAATGGGACGAATCTCTCTTCTCCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATGCAGCAGACC TTCATCTCTTCAGCACAAAGAAGTCACTGCTGCTGGATGAGACC

		CTCCTAGAAAAATTTCCACTGAACCTTACCGCAACTGAATGAACGTG GAAGCATGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCTGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAACACT CTTTATCTGACAGAGAAGAAGTATAGCCCTGTTCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:17	6CG3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAAGAGGGCCATGATG CTCCTGGCACAAATGGGAAGAACCTCTCCTTCTGCCTGTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAGGCTCAAGCCATCTTGTCTCCATGAGATGATCCAGCAGACC TTCAATTCTTCAGCACAAAGGACTCATCTGCTGCTGGGAACAGAGC CTCCTAGAAAAATTTCCACTGAACCTAACCGCAGCTGAATGACCTG GAAGCCTGCGTGTGATACAGGAAGTGGGTGGAAGAGACTCCCTGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAACACT CTTTATCTGACAGAGAAGAATACAGCCCTGTCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:18	3CG7	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGTAGGGCCTTGATG CTCCTGGCACAAATGGGAAGAACATCTCCCCTTCTGCCTGTGAAGGAC AGACATGATTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGCCATCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCATCTGCTGCTGGGAACAGAAC CTCCTAGAAAAATTTCCACTGAACCTTACCGCAACTGAATAACCTG GAAGCATGTGTGATACAGGAGGTTGGATGGAAGAGACTCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGGAAGTACTTCAAAGAACACT CTTTATCTAATAGAGAGGAAATACAGCCCTGTCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:19	1D3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGCACAAATGGGAAGAACATCTCATTCTCCTGCCTGTGAAGGAC AGACATGATTCGGATTCCCCCAGGAGGAGTTGATGGCCACCAGTTC CAGAAGACTCAAGCCATCTGCTCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTGCTGGGAACAGAGC CTCCTAGAAAAATTTCCACTGAACCTTACCGCAACTGAATGACCTG GAAGCATGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCTGATG AATGAGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAACACT CTTTATCTGATGGAGAAGAAATACAGCCCTGTCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:20	2G4	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCATGATG CTCCTGGCACAAATGAGCAGAACATCTCCTCCTCCTGTGTGATGGAC AGACATGACTTGAATTTCCCCCAGGAGGAATTGATGATAAACAGTTC CAGAAGGCTCCAGCCATCTGCTCCTCCATGAGGTGATTGAGCAGACC TTCAATCTCTTCAGCACAGAGGACTCATCTGCTGCTGGGAACAGACC CTCCTAGAAAAATTTCCACTGAACCTTACCGCAACTGAATGACCTG GAAGCATGTGTGATGAGGAGGAGGGTGGGAGAAACTCCCTGATG AATGCGGACTCCATCTGGCTGTGAGGAATACTTCAAAGAACACT CTTTATCTGACAAAGAAGAAGTATAGCCCTGTTCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA

		AGATTAAGGAGGAAGGAA
SEQ ID NO:21	1A1	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCTCATTTCCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGTGTTGATGGCAACCAGTTC CAGAAGGCCAACGCCATCTGCCTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAGAGGACTCATCTGCTGCTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTCACCAGCACTGAATGACCTG GAAGCCTGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGGAAATACTTCAAAGAACATCACT CTTATCTAATGGAGAAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:22	ID10	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCTCATTTCCTGCCTGAAGGAC AGACATGATTCGGATTCCCCCAGGAGGAGTTGATGGCCACCAGTTC CAGAAGACTCAAGGCCATCTCTGCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTGCTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTACCAAGCACTGAATGACCTG GAAGCATGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGGAAATACTTCAAAGAACATCACT CTTATCTGATGGAGAAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:23	1F6	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTGTGATG ATAATGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGGCCATCTCTGCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTAACCAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGCTGGGTGGAAGAGACTCCCCTGATG AATGTGAGCTCCATCCTGGCTGTGAGGAAATACTTCAAAGAACATCACT CTTATCTAACAGAGAAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:24	2A10	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCTCATTTCCTGCCTGAAGGAC AGATATGATTCGGATTCCCCCAGGAGGTGTTGATGGCAACCAGTTC CAGAAGGCTCAAGGCCATCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGAACAGAGC CTCCTAGAAAAAATTTCCACTGAACTTACCAAGCACTGAATAACCTG GAAGCATGTGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCTGATG AATGAGGACTCCATCCTGGCTGTGAGGAAATACTTCAAAGAACATCACT CTTATCTGATGGAGAAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:25	2C3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACATGACTTGGATTCCCTCAGGAGGAGTTGATGGCAACCAGTCC

		CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGATACTTGGGATGCCGACC CTTTAGAAAAATTTCACACTGAACCTAACAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACCCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAACTTCAAAGAATCACT CTTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:26	2D1	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGGAGGGCTTGATA CTCCTGGCACAATGGGACGAATCTCCTCTCCTGCCTGAAGGAC AGACAAGACTTTGGATTCCCCCAGGAGGAGTTGATGGCAACCGGTT CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTGCTTGGGAAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATAACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAACTTCAAAGAATCACT CTCTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:27	2D10	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGGAGGGCTTGATA CTCCTGGCACAATGGGAGAGTCTCCTCTCCTGCCTGAAGGAC AGACATGACTTTGGATTCCCCCAGGAGGAGTTGATGGCAACCGAGTT CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGAAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATAACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAACTTCAAAGAATCACT CTCTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:28	2D7	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGGAGGGCTTGATA CTCCTGGCACAATGGGAGAGAATCTCCTCTCCTGCCTGAAGGAC AGACATGACTTCAGATTCCCCCAGGAGGAGTTGATGGCAACCGAGTT CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGAAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTTACCAAGCAACTGAATAACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCCTGATG AATGTGGACTCTATCCTGGCTGTGAAGAAACTTCAAAGAATCACT CTTATCTGACAGAGAGGAAATACAGCCCTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:29	2D9	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGGAGGGCTTGATA CTCCTGGCACAATGGGAGAGAATCTCCTCTCCTGCCTGAAGGAC AGACATGACTTTGGATTCCCCCAGGAGGAGTTGATGGCAACCGAGTT CAGAAGGCTCAAGCCATCTCTGCCTCCATGAGATGATCCAGCAGACT TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGAAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTAACAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGAGACTCCCCCTGGT AATGTGGACTCCATCCTGGCTGTGAAGAAACTTCAAAGAATCACT

		CTTTATCTGACAGAGAAGAAATACAGCCCTTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:30	2DA2	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGGGAAGAAATCTCTCCTTCTCCTGCCTGAAGGAC AGACAGGACTTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAAAGAACCTCATCTGTGCCTGGGAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTCCACCAGCAACTGAATGAACCTG GAAGCATGTGTGATACAGGAGGTTGGGGTGGAAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAATACTTCAAAGAACACT CTTTATCTAATAGAGAGGAATACAGCCCTTGTGCATGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:31	2DH9	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCCTGATG CTCCTGGCACAAATGGACGAATCTCTCCTTCTCCTGCCTGAAGGAC AGATATGATTTCGGATTCCCCCAGGGGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTGCCTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAAAGGATTCTACATCTGTGCCTGGGAACAGAGC CTCCTAGAAAAATTTCACACTGAACCTCACCGGAGCTGAATGACCTG GAAGCCTGTGTGATACAGGAGGTTGGGGTGGAAAGAGACCCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGGAAGTACTTCAAAGAACACT CTTTATCTGACAGAGAACGATAGCCCTTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:32	2G11	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGCCCTGATA CTCCTGGCACAAATGGGAAGAAATCTCTCCTTCTCCTGCCTGAAGGAC AGACATGACTTGGACTTCCCAGGAGGAGTTGATGGCAACCAGTC CAGAAGACTCAAGCCATCTGTGCCTCCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGTGATACTTGGGAACAGAGC CTCCTAGAAAAATTCTACATTGAACCTTCCAGCAGCTGAATGACCTG GAAGCCTGCCTGTGATACAGGAGGTTGGGGTGGAAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAGAAAATACTTCAAAGAACACT CTTTATCTGACAGAGGAGAAATACAGCCCTTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:33	2G12	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTGATG CTCATGGCACAAATGAGGAGAATCTCTCCTTCCCCCGCCTGAAGGAC AGATATGATTTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCTATCTCCTTTCCATGAGATGATGCAGCAGACC TTCAATCTCTTCAGCACAAAGAACCTCATCTGTGCCTGGGATGAGACC CTCCTAGACAAATTCTACACTGAACCTCACAGCAGCTGAATGACTTG GAAGCCTGTGTGATGCAGGAGGGAGGGTGGGAGAAACTCCCCTGATG AATGCGGACTCCATCTGGCTGTGAAGAAATACTTCCGAAGAACACT CTCTATCTGACAGAGAACATACAGCCCTTGTGCCTGGGAGGCTGTC AGAGCAGAAATCATGAGATCTTCTTTCAACAAACTTGCAAAA AGATTAAGGAGGAAGGAA

SEQ ID NO:34	2H9	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGATA CTCCTGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTCTCATGAGATGATCCAGCAGACC TTCAATCTCTCAGCACAAAGGACTCATCTGCTACTTGGAACAGAGC CTCCTAGAAAATTTCACTGAACCTAACAGCAGCTGAATGACCTA GAAGCCTGTGTGACACAGGAGGTGGGGTGGAAAGAGACTCCCCTGATG AATGAGGACTCTATCCTGGCTGTGAAGAAATACTCCAAGAACACT CTTATCTGACAGAGAACATGAGATCTTCTCTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:35	6BC11	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTGATA CTCCTGGCACAAATGGGAAGAATCTCTCCTTCTGCCTGAAGGAC AGATATGATTTCGGATTCCCCCAGGAGGAGTTGATGGCAACCAGCTC CAGAAGGCTCAAGCCATCTGTCTCATGAGATGATCCAGCAGACC TTCAATCTCTCAGCACAAAGGATTCATCTGCTGCTTGGAACAGAGC CTCCTAGAAAATTTCACTGAACCTAACAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTGGAGTGGAAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAATACTCCAAGAACACT CTTATCTGACAGAGAACATGAGATCTTCTCTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:36	2DH12	CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGPQEEDGNQF QKAQAIISVLHEMIQQTFNLFKDSSAAWEQTLLEKFSTELYQQLNDL EACVIQEVGVKETPLMNVDSDILAVRKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:37	2CA3	CDLPQTHSLGDRRAMILLAQMGRISPFSCLKDRYDFGPQEEDGNQF QKAQAIISVLHEMIQQTFNLFKDSSAAWEQSLLEKFSTELYQQLNEL EACVIQEVGVGETPLMNGDSILAVKKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:38	4AB9	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGPREEFDGNQF QKAQAIISVLHEMMQQTFNLFKDSSAAWDETLLEKFSTELYQQLNEL EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYSPCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:39	2DA4	CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGPQEEDDSNQF QKAQAIISVLHEMMQQTFNLFKDSSAAWDETLLEKFSTELYQQLNDL EACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:40	3DA11	CDLPQTHSLGNRRALVLLAQMGRISPFSCLKDRYDFGPQEEDGNQF QKAQAIISVLHEMIQQTFNLFKDSSAAWDETLLEKFSTELYQQLNDL EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:41	2DB11	CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGPQEEDGNQF QKAQAIISVLHEMIQQTFNLFKDSSAAWDETLLEKFSTELYQQLNDL EACVIQEVGVEETPLMNVDSDILAVRKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:42	2CA5	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRQDFGPQEEDGNRF QKAQAIISVLHEMIQQTFNLFKDSSAAWEQSLLEKFSTELYQQLNDL

		EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLIERKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:43	2G6	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKDSSATWEQSLLEKFSTELNQQLNDL EACVIQEVGVEETPLMNVD PILAVKKYFQRITLYLTEKKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:44	3AH7	CDLPQTHSLGNRRALILLAQMRRISPFSCLKDRHDFGPQEEFDNSQF QKAQAI SVLHEMIQQTFLFSTKDSSAAWEQSLLEKFSTELHQQLNEL EACV VQEVGVEETPLMNEDSILAVKKYLQQRITLYLTEKKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:45	2G5	CDLPQTHSLGNRRALM LLAQMGRISPFSCLKDRQDFGPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKDSSAAWEQSLLEKFSTELYQQLNDL EACVIQEVGVEETPLMNVD SILAVRKYFQRITLYLIERKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:46	2BA8	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKDSSAAWEQSLLEKFSTELYQQLNDL EACVIQEVGVEETPLMNVD SILAVRKYFQRITLYLIERKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:47	1F3	CDLPQTHSLGNRRALILLGQMGRISHFSCLKDRHDFGPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKDSSVAWDERLLDKLYT ELYQQLNDL EACVMQEWWVG GTPLMNEDSILAVRKYFQRITLYLTEKKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:48	4BE10	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGPQEEFDGNQF QKAQAI SVLHEIMQQTFLFSTKNSSAWDETLLKEFSTELYQQLNEL EACVIQGVGVEETPLMNEDSILAVRKYFQRITLYLTEKKYS PCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:49	2DD9	CDLPQTHSLGNRRALM LLAQMGRISPFSCLKDRYDFGPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKDSSAAWEQSLLEKFSTG LYQQLNDL EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYS PCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:50	3CA1	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGLPQEEFDGNQF QKAQAI SVLHEMIQQTFLFSTKNSSAAWDETLLKEFSTELYQQLNNL EACVIQEVGMEETPLMNVD SILAVKKYFQRITLYLTEKKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:51	2F8	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGPQEEFDGNQF QKAQAI SVLHEMMQQTFLFSTKNSSAAWDETLLKEFSTELYQQLNEL EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYS PCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:52	6CG3	CDLPQTHSLGNKRAMMLAQMGRISPFSCLKDRHDFGPQEEFDGNQF QRAQAIFVLHEMIQQTFLFFSTKDSSAAWEQSLLEKFSTELNQQLNDL EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:53	3CG7	CDLPQTHSLGNSRALM LLAQMGRISPFSCLKDRHDFGPQEEFDGNQF QKAQAI SAHEMIQQTFLFSTKDSSAAWEQNLLEKFSTELYQQLNNL EACVIQEVGMEETPLMNVD SILAVRKYFQRITLYLIERKYS PCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:54	1D3	CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRHDFGPQEEFDGHQF QKTQAISVLHEMIQQTFLFSTKDSSAAWEQSLLEKFSTELYQQLNDL

		EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLMEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:55	2G4	CDLPQTHSLGNRRAMMLAQMSRISPSSCLMDRHDFFPQEEFDKQF QKAQAIISVLHEVIQQTFLNSTEDSSAAWEQTLLEKFSTELYQQLNNDL EACVMQEERVGGETPLMNADSI LAVRKYFQRITLYLT KKKYSPCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:56	1A1	CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRYDFGF P QEVFDGNQF QKAQAIISAFHEMMQQTFLNSTEDSSAAWEQS LLEKFSTELHQQLNDL EACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLMEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:57	1D10	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFRF P QEEFDGNQL QKTQAISVLHEMIQQTFLNSTKDSSATWEQS LLEKFSTELNQQLNDL EACVIQGVGVEETPPMVDSI LAVKKYFQRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:58	1F6	CDLPQTHSLGNRRALIMAQMGRISPFSCLKDRHDFGF P QEEFDGNQF QKAQAIISVLHEMIQQTFLNSTKDSSATWEQS LLEKFSTELNQQLNDL EACVIQEA GVEETPLMNVDSI LAVKKYFQRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:59	2A10	CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRYDFGF P QEVFDGNQF QKAQAIISAFHEMIQQTFLNSTKDSSATWEQS LLEKFSTELYQQLNNDL EACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLMEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:60	2C3	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGF P QEEFDGNQF QKAQAIISVLHEMIQQTFLNSTKDSSDTWDATLEKFSTELNQQLNDL EACVIQEVGVEETPLMNVDSI LAVKKYFQRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:61	2D1	CDLPQTHSLGNRRALILLAQMRRISPFSCLKDRHDFGF P QEEFDGNQF QKAQAIISAFHEMIQQTFLNSTKDSSATWEQS LLEKFSTELYQQLNNDL EACVIQEVGMEETPLMNEDSILAVKKYFQRITLYLTERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:62	2D10	CDLPQTHSLGNRRALILLAQMGRVSPFSCLKDRHDFGF P QEEFDGNQF QKAQAIISAFHEMIQQTFLNSTKDSSATWEQS LLEKFSTELYQQLNNDL EACVIQEVGVEETPLMNVDSI LAVKKYFRRITLYLTERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:63	2D7	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFRF P QEEFDGNQF QKAQAIISVLHEMIQQTFLNSTKDSSATWEQS LLEKFSTELYQQLNNDL EACVIQEVGVEETPLMNVDSI LAVKKYFQRITLYLTERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:64	2D9	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGF P QEEFDGNQF QKAQAIISVLHEMIQQTFLNSTKDSSATWEQS LLEKFSTELNQQLNDL EACVIQEVGVEETPLVNVD S I LAVKKYFQRITLYLTERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:65	2DA2	CDLPQTHSLGNRRPLILLAQMGRISPFSCLKDRQDFGPQEEFDGNQF QKAQAIISVLHEMMQQTFLNSTKNSSAAWEQS LLEKFSTELHQQLNEL EACVIQEVGVEETPLMNVDSI LAVKKYFQRITLYLIERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:66	2DH9	CDLPQTHSPGNRRALILLAQMGRISPFSCLKDRYDFGF P QGEFDGNQF QKAQAIISVLHEMMQQTFLNSTKDSSAAWEQS LLEKFSTELYRQLNDL

		EACVIQEVGVEETPLMNVDISILAVRKYFQRITLYLTEKKHSPCSWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:67	2G11	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGLPQEEFDGNQF QKTQAIISVLHEMIQQTFNLFSTKDSSDTWEQSLEKFYIELFQQLNDL EACVIQEVGVEETPLMNVDISILAVRKYFQRITLYLTEEKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:68	2G12	CDLPQTHSLGNRRTLMLMAQMRRISPFPRLKDRYDFGFPQEVFDGNQF QKAQAIIFLFHEMMQQTFNLFSTKNSSAAWDETLLDKFYTELQQLNDL EACVMQEGRVGETPLMNADSI LAVK KYFRRITLYLTEKKYSPCAWEAV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:69	2H9	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQEEFDGNQF QKAQAIISVLHEMIQQTFNLFSTKDSSATWEQSLEKFSTELNQQLNDL EACVTQEVGVEETPLMNEDSI LAVK KYFQRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:70	6BC11	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGFPQEEFDGNQL QKAQAIISVLHEMIQQTFNLFSTKDSSAAWEQSLEKFSTELNQQLNDL EACVIQEVGVEETPLMNVDISILAVK KYFQRITLYLTERKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:71	t19bb	CDLPQTHSLGXXRAXXLLXQMXRXSXF SCLKDRXDFGXPXEEDXXXF QXXQAIXXXHEXXQQTFNXFSTKXSSXXWXXXLXKXXTXLXQQLNXL EACVXQXVXXXXTPLMNDXI LAVXK KYXQRITLYLXEXKYSPCXWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:72	CH1.1	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCCTTCTCCTGCTGATGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGACAACCAGTTC CAGAAGGCTCAAGGCATCTGTCTCCATGAGATGATCCAACAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGATGAGACA CTTCTAGACAAATTCTACACTGAACTTACCAAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCTGATG AATGAGGACTCCATCTGGCTGTGAAGAAATACTCCGAAGAACACT CTCTATCTGACAGAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:73	CH1.2	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGGCCTTGATA CTCCTGGCACAAATGGGAAGAATCTCCTTCTCCTGCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTTC CAGAAGGCTCAAGGCATCTGTCTCCATGAGATGATCCAGCAGACC TTCCATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGAACAGAGC CTCCTAGAAAATTTCACACTGAACTAACCAAGCAGCTGAATGACCTG GAAGCCTGCGTGATACAGGAGGTTGGGTGGAAGAGACTCCCCTGATG AATGTGGACTCCATCCTGGCTGTGAAGAAATACTCCGAAGAACACT CTTATCTGACAGAGAAATACAGCCCTTGTGCCTGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:74	CH1.3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTGTGATG ATAATGGCACAAATGGGAAGAATCTCCTTCTCCTGCTGAAGGAC AGACATGACTTGGATTCCCTCAGGAGGAGTTGATGGCAACCAGTTC

		CAGAAGGCTCAAGCCATCTGTCCCTCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGATGAGACA CTTCTAGACAAATTCTACACTGAACCTTACCGAGCTGAATGACCTG GAAGCCTGTATGATGCAGGAGGTTGGAGTGGAAAGACACTCCTCTGATG AATGTGGACTCTATCCTGACTGTGAGAAAATACCTTCGAAGAACACT CTTTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:75	CH1.4	TGTGATCTGCCTCAGACCCACAGCCTGGTAATAGGAGGGCCTGATA CTCCTGGCACAAATGGGAAGAAATCTCTCCTTCCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGGTGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTCCCTCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAGAGGACTCATCTGCTGCTTGGGATGAGACC CTCCTAGACAAATTCTACATTGAACCTTCCAGCAACTGAATGACCTG GAAGCCTGTGTATGATGCAGGAGGAGAGGGTGGGAGAAACTCCCCTGATG AATGCGGACTCCATTTGGCTGTGAAGAAAATACCTCCAAAGAACACT CTTTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:76	CH2.1	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTTGATG ATAATGGCACAAATGGGAAGAAATCTCTCCTTCCTGCCTGAAGGAC AGACATGACTTGGATTTCCTCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTCCCTCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGATGAGACA CTTCTAGACAAATTCTACACTGAACCTTACCGAGCTGAATGACCTG GAAGCCTGTATGATACAGGAGGTTGGGGTGGAAAGAGACTCCCCTGATG AATGAGGACTCCATTTGGCTGTGAAGAAAATACCTCCGAAGAACACT CTCTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:77	CH2.2	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTTGATG CTCCTGGCACAAATGGGAAGAAATCTCTCCTTCCTGCCTGAAGGAC AGACATGACTTGGATTCCCCCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTCCCTCATGAGATGATCCAACAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGATGAGACA CTTCTAGACAAATTCTACACTGAACCTTACCGAGCTGAATGACCTG GAAGCCTGTATGATGCAGGAGGTTGGAGTGGAAAGACACTCCTCTGATG AATGTGGACTCTATCCTGACTGTGAAGAAAATACCTCCGAAGAACACT CTTTATCTGACAGAGAAGAAATACAGCCCTGTGCCTGGGAGGTTGTC AGAGCAGAAATCATGAGATCTTCTCTTTCAACAAACTTGCAAAAA AGATTAAGGAGGAAGGAA
SEQ ID NO:78	CH2.3	TGTGATCTGCCTCAGACCCACAGCCTGGTAACAGGAGGACTTTGATG ATAATGGCACAAATGGGAAGAAATCTCTCCTTCCTGCCTGAAGGAC AGACATGACTTGGATTTCCTCAGGAGGAGTTGATGGCAACCAGTC CAGAAGGCTCAAGCCATCTGTCCCTCATGAGATGATCCAGCAGACC TTCAATCTCTTCAGCACAAAGGACTCATCTGCTACTTGGGATGAGACA CTTCTAGACAAATTCTACACTGAACCTTACCGAGCTGAATGACCTG GAAGCCTGTATGATGCAGGAGGTTGGAGTGGAAAGACACTCCTCTGATG

		AATGAGGACTCCATCTTGGCTGTGAAGAAATACTTCCGAAGAATCACTCTCTATCTGACAGAGAAGAAATACAGCCCCTGTGCCTGGGAGGTTGTCAGAGCAGAAATCATGAGATCTTCTCTCAACAAACTTGCAAAAAAGATTAAGGAGGAAGGAA
SEQ ID NO:79	CH1.1	CDLPQTHSLGNRRALILLAQMGRISPFSCLMDRHDGFHQEEFDNNQFQKAQAIISVLHEMIQQTFNLFSTKDSSATWDETLDDKFYTELYQQLNDL EACVIQEVGVEETPLMNEDSILAVKKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:80	CH1.2	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDGFHQEEFDGNQFQKAQGIISVLHEMIQQTFHLFSTKDSSATWEQSLLEKFSTELNQQLNDL EACVIQEVGVEETPLMNVDISILAVKKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:81	CH1.3	CDLPQTHSLGNRRTLMIMAQMGRISPFSCLKDRHDGFHQEEFDGNQFQKAQAIISVLHEMIQQTFNLFSTKDSSATWDETLDDKFYTELYQQLNDL EACMMQEVGVEDTPLMNVDISILTVRKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:82	CH1.4	CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDGFHQEEFGGNQFQKAQAIISVLHEMIQQTFNLFSTEDSSAAWDETLDDKFYIELFQQLNDL EACVMQEERVGETPLMNADSILAVKKYFQRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:83	CH2.1	CDLPQTHSLGNRRTLMIMAQMGRISPFSCLKDRHDGFHQEEFDGNQFQKAQAIISVLHEMIQQTFNLFSTKDSSATWDETLDDKFYTELYQQLNDL EACMIQEVGVEETPLMNEDSILAVKKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:84	CH2.2	CDLPQTHSLGNRRALILLAQMGRISPFSCLMDRHDGFHQEEFDNNQFQKAQAIISVLHEMIQQTFNLFSTKDSSATWDETLDDKFYTELYQQLNDL EACMMQEVGVEETPLMNVDISILTVKKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE
SEQ ID NO:85	CH2.3	CDLPQTHSLGNRRTLMIMAQMGRISPFSCLKDRHDGFHQEEFDGNQFQKAQAIISVLHEMIQQTFNLFSTKDSSATWDETLDDKFYTELYQQLNDL EACMMQEVGVEETPLMNEDSILAVKKYFRRITLYLTEKKYSPCAWEVV RAEIMRSFSFSTNLQKRLRRKE

**WHAT IS CLAIMED IS:**

1. An isolated or recombinant nucleic acid, comprising:
  - a polynucleotide sequence selected from the group consisting of:
    - (a) SEQ ID NO:1 to SEQ ID NO:35, or a complementary polynucleotide sequence thereof;
    - (b) a polynucleotide sequence encoding a polypeptide selected from SEQ ID NO:36 to SEQ ID NO:70, or a complementary polynucleotide sequence thereof;
    - (c) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide sequence (a) or (b); and
    - (d) a polynucleotide sequence comprising a fragment of (a), (b), or (c), which fragment encodes a polypeptide having antiproliferative activity in a human Daudi cell line - based assay.
2. An isolated or recombinant nucleic acid, comprising:
  - a polynucleotide sequence selected from the group consisting of:
    - (a) SEQ ID NO:72 to SEQ ID NO:78, or a complementary polynucleotide sequence thereof;
    - (b) a polynucleotide sequence encoding a polypeptide selected from SEQ ID NO:79 to SEQ ID NO:85, or a complementary polynucleotide sequence thereof;
    - (c) a polynucleotide sequence which hybridizes under highly stringent conditions over substantially the entire length of polynucleotide sequence (a) or (b); and
    - (d) a polynucleotide sequence comprising a fragment of (a), (b) or (c), which fragment encodes a polypeptide having antiviral activity in a murine cell line/EMCV - based assay.
3. An isolated or recombinant nucleic acid, comprising:
  - a polynucleotide sequence encoding a polypeptide, the polypeptide comprising the amino acid sequence: CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QQLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-

WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof, where  $X_{11}$  is N or D;  $X_{12}$  is R, S, or K;  $X_{15}$  is L or M;  $X_{16}$  is I, M, or V;  $X_{19}$  is A or G;  $X_{22}$  is G or R;  $X_{24}$  is I or T;  $X_{26}$  is P or H;  $X_{34}$  is H, Y or Q;  $X_{38}$  is F or L;  $X_{40}$  is Q or R;  $X_{45}$  is G or S;  $X_{46}$  is N or H;  $X_{47}$  is Q or R;  $X_{50}$  is K or R;  $X_{51}$  is A or T;  $X_{55}$  is S or F;  $X_{56}$  5 is V or A;  $X_{57}$  is L or F;  $X_{60}$  is M or I;  $X_{61}$  is I or M;  $X_{67}$  is L or F;  $X_{72}$  is D or N;  $X_{75}$  is A or V;  $X_{76}$  is A or T;  $X_{78}$  is E or D;  $X_{79}$  is Q or E;  $X_{80}$  is S, R, T, or N;  $X_{83}$  is E or D;  $X_{85}$  is F or L;  $X_{86}$  is S or Y;  $X_{88}$  is E or G;  $X_{90}$  is Y, H, N;  $X_{95}$  is D, E, or N;  $X_{101}$  is I, M, or V;  $X_{103}$  is E or G;  $X_{105}$  is G or W;  $X_{106}$  is V or M;  $X_{107}$  is E, G, or K;  $X_{108}$  is E or G;  $X_{114}$  is V, 10 E, or G;  $X_{116}$  is S or P;  $X_{121}$  is K or R;  $X_{124}$  is F or L;  $X_{132}$  is T, I, or M;  $X_{134}$  is K or R; and  $X_{140}$  is A or S.

4. The nucleic acid of claim 3, said polypeptide having antiproliferative activity in a human Daudi cell line-based cell proliferation assay or antiviral activity in a human WISH cell/EMCV-based assay.

5. The nucleic acid of claim 3, wherein the encoded polypeptide has an 15 antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in a human Daudi cell line - based assay or an antiviral activity of at least about  $2.1 \times 10^7$  units/milligram in a human WISH cell/EMCV-based assay.

6. The nucleic acid of claim 3, wherein the encoded polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:36 20 to SEQ ID NO:54.

7. The nucleic acid of claim 3, said nucleic acid comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO:1 to SEQ ID NO:19.

8. An isolated or recombinant nucleic acid comprising a polynucleotide 25 sequence encoding a polypeptide, the polypeptide comprising:

an amino acid sequence comprising at least 20 contiguous amino acids of any one of SEQ ID NOS:36-70, and one or more of amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, wherein the numbering of the amino acids corresponds to that of SEQ ID NO:36.

9. The nucleic acid of claim 8, wherein the encoded polypeptide is 166 amino acids in length.

10. The nucleic acid of claim 8, wherein the encoded polypeptide has an antiproliferative activity in a human Daudi cell line - based assay.

5 11. The nucleic acid of claim 8, wherein the encoded polypeptide has an antiviral activity in a human WISH cell/EMCV-based assay.

12. The nucleic acid of claim 8, wherein the encoded polypeptide comprises amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166.

10 13. The nucleic acid of claim 8, wherein the encoded polypeptide comprises at least 50 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

14. The nucleic acid of claim 8, wherein the encoded polypeptide comprises at least 100 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

15 15. The nucleic acid of claim 8, wherein the encoded polypeptide comprises at least 150 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

16. The nucleic acid of claim 8, wherein the encoded polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:45, and SEQ ID NO:46.

20 17. The nucleic acid of claim 8, comprising a polynucleotide sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:10, and SEQ ID NO:11.

18. An isolated or recombinant nucleic acid comprising a polynucleotide sequence encoding a polypeptide, the polypeptide comprising:

25 an amino acid sequence comprising at least 155 contiguous amino acids of any one of SEQ ID NOS:36-70, said amino acid sequence comprising amino acids Lys160

and Glu166, wherein the numbering of the amino acids corresponds to that of SEQ ID NO:36.

19. The nucleic acid of claim 18, wherein the encoded polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:36, 5 SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:45, and SEQ ID NO:46.
20. A cell comprising the nucleic acid of claim 1, 2, 8, or 18.
21. The cell of claim 20, wherein the cell expresses a polypeptide encoded by the nucleic acid.
- 10 22. A vector comprising the nucleic acid of claim 1, 2, 8, or 18.
23. The vector of claim 20, wherein the vector comprises a plasmid, a cosmid, a phage, or a virus.
24. The vector of claim 22, wherein the vector is an expression vector.
25. A cell transduced by the vector of claim 22.
- 15 26. A composition comprising the nucleic acid of claim 1, 2, 8, or 18, and an excipient.
27. The composition of claim 26, wherein the excipient is a pharmaceutically acceptable excipient.
- 20 28. A composition produced by digesting one or more nucleic acids of claim 1, 2, 3, 8, or 18 with a restriction endonuclease, an RNase, or a DNase.
29. A composition produced by a process comprising incubating one or more nucleic acids of claim 1, 2, 3, 8, or 18 in the presence of deoxyribonucleotide triphosphates and a nucleic acid polymerase.
- 25 30. The composition of claim 29, wherein the nucleic acid polymerase is a thermostable polymerase.

31. An isolated or recombinant polypeptide encoded by the nucleic acid of acid claim 1, 2, 3, 8, or 18.

32. The isolated or recombinant polypeptide of claim 31, comprising a sequence selected from the group consisting of: SEQ ID NO:36 to SEQ ID NO:70 or SEQ 5 ID NO:79 to SEQ ID NO:85.

33. The polypeptide of claim 31, having an antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram (mg) in a human Daudi cell line - based assay or an antiviral activity of at least about  $2.1 \times 10^7$  units/milligram in a human WISH cell/EMCV-based assay.

10 34. An isolated or recombinant polypeptide, comprising:  
the amino acid sequence: CDLPQTHSLG-X<sub>11</sub>-X<sub>12</sub>-RA-X<sub>15</sub>-X<sub>16</sub>-LL-X<sub>19</sub>-QM-X<sub>22</sub>-R-X<sub>24</sub>-S-X<sub>26</sub>-FSCLKDR-X<sub>34</sub>-DFG-X<sub>38</sub>-P-X<sub>40</sub>-EEFD-X<sub>45</sub>-X<sub>46</sub>-X<sub>47</sub>-FQ-X<sub>50</sub>-X<sub>51</sub>-QAI-X<sub>55</sub>-X<sub>56</sub>-X<sub>57</sub>-HE-X<sub>60</sub>-X<sub>61</sub>-QQTFN-X<sub>67</sub>-FSTK-X<sub>72</sub>-SS-X<sub>75</sub>-X<sub>76</sub>-W-X<sub>78</sub>-X<sub>79</sub>-X<sub>80</sub>-LL-X<sub>83</sub>-K-X<sub>85</sub>-X<sub>86</sub>-T-X<sub>88</sub>-L-X<sub>90</sub>-QQLN-X<sub>95</sub>-LEACV-X<sub>101</sub>-Q-X<sub>103</sub>-V-X<sub>105</sub>-X<sub>106</sub>-X<sub>107</sub>-X<sub>108</sub>-TPLMN-X<sub>114</sub>-15 D-X<sub>116</sub>-ILAV-X<sub>121</sub>-KY-X<sub>124</sub>-QRITLYL-X<sub>132</sub>-E-X<sub>134</sub>-KYSPC-X<sub>140</sub>-WEVVRAEIMRSFSFSTNLQKRLRRKE, or a conservatively substituted variation thereof;

wherein X<sub>11</sub> is N or D; X<sub>12</sub> is R, S, or K; X<sub>15</sub> is L or M; X<sub>16</sub> is I, M, or V; X<sub>19</sub> is A or G; X<sub>22</sub> is G or R; X<sub>24</sub> is I or T; X<sub>26</sub> is P or H; X<sub>34</sub> is H, Y or Q; X<sub>38</sub> is F or L; X<sub>40</sub> is Q or R; X<sub>45</sub> is G or S; X<sub>46</sub> is N or H; X<sub>47</sub> is Q or R; X<sub>50</sub> is K or R; X<sub>51</sub> is A or T; X<sub>55</sub> is S or F; X<sub>56</sub> is V or A; X<sub>57</sub> is L or F; X<sub>60</sub> is M or I; X<sub>61</sub> is I or M; X<sub>67</sub> is L or F; X<sub>72</sub> is D or N; X<sub>75</sub> is A or V; X<sub>76</sub> is A or T; X<sub>78</sub> is E or D; X<sub>79</sub> is Q or E; X<sub>80</sub> is S, R, T, or N; X<sub>83</sub> is E or D; X<sub>85</sub> is F or L; X<sub>86</sub> is S or Y; X<sub>88</sub> is E or G; X<sub>90</sub> is Y, H, N; X<sub>95</sub> is D, E, or N; X<sub>101</sub> is I, M, or V; X<sub>103</sub> is E or G; X<sub>105</sub> is G or W; X<sub>106</sub> is V or M; X<sub>107</sub> is E, G, or K; X<sub>108</sub> is E or G; 25 X<sub>114</sub> is V, E, or G; X<sub>116</sub> is S or P; X<sub>121</sub> is K or R; X<sub>124</sub> is F or L; X<sub>132</sub> is T, I, or M; X<sub>134</sub> is K or R; and X<sub>140</sub> is A or S.

30 35. The polypeptide of claim 34, having antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in a human Daudi cell line - based assay or antiviral activity of at least about  $2.1 \times 10^7$  units/milligram in a human WISH cell/EMCV-based assay.

36. The polypeptide of claim 34, comprising a sequence selected from the group consisting of: SEQ ID NO:36 to SEQ ID NO:54.

37. A polypeptide comprising at least 100 contiguous amino acids of a protein encoded by a coding polynucleotide sequence, the polynucleotide sequence 5 selected from the group consisting of:

- (a) SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78;
- (b) a coding polynucleotide sequence that encodes a first polypeptide selected from SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85; and
- 10 (c) a complementary polynucleotide sequence which hybridizes under highly stringent conditions over substantially an entire length of a polynucleotide sequence of (a) or (b).

38. The polypeptide of claim 37, said polypeptide having an antiproliferative activity in a human Daudi cell line-based cell proliferation assay or an 15 antiviral activity in a human WISH cell/EMCV-based assay.

39. The polypeptide of claim 37, wherein the polypeptide specifically binds to a human alpha-interferon receptor.

40. The polypeptide of claim 37, comprising at least 150 contiguous amino acids of the encoded protein.

20 41. An isolated or recombinant polypeptide, comprising: an amino acid sequence comprising at least 50 contiguous amino acids of any one of SEQ ID NOS:36-70, the amino acid sequence comprising one or more of amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, wherein the numbering of the amino acids corresponds to that of SEQ ID NO:36.

25 42. The polypeptide of claim 41, wherein the polypeptide binds a human alpha-interferon receptor.

**43.** The polypeptide of claim 41, said polypeptide exhibiting an antiproliferative activity in a human Daudi cell line-based cell proliferation assay or an antiviral activity in a human WISH cell/EMCV-based assay.

5       **44.** The polypeptide of claim 41, having an antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in a human Daudi cell line - based assay or an antiviral activity of at least about  $2.1 \times 10^7$  units/milligram in a human WISH cell/EMCV-based assay.

**45.** The polypeptide of claim 41, wherein the polypeptide is 166 amino acids in length.

10       **46.** The polypeptide of claim 41, said polypeptide comprising amino acids Ala19, (Tyr or Gln)34, Gly37, Phe38, Lys71, Ala76, Tyr90, Ile132, Arg134, Phe152, Lys160, and Glu166, wherein the numbering of the amino acids of said polypeptide corresponds to the numbering of amino acids in SEQ ID NO:36.

15       **47.** The polypeptide of claim 41, comprising at least 100 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

**48.** The polypeptide of claim 41, comprising at least 150 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

**49.** The polypeptide of claim 41, comprising at least 155 contiguous amino acid residues of any one of SEQ ID NOS:36-70.

20       **50.** The polypeptide of claim 41, comprising an amino acid sequence selected from the group consisting of: SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:45, and SEQ ID NO:46.

25       **51.** An isolated or recombinant polypeptide comprising an amino acid sequence comprising at least 155 contiguous amino acids of any one of SEQ ID NOS:36-70, the isolated or recombinant polypeptide comprising amino acids Lys160 and Glu166, wherein the numbering of the amino acids corresponds to that of SEQ ID NO:36.

**52.** The polypeptide of claim 51, comprising an amino acid sequence selected from the group consisting of: SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:45, and SEQ ID NO:46.

**53.** The polypeptide of claim 51, said polypeptide having an antiproliferative activity of at least about  $8.3 \times 10^6$  units/milligram in milligram in a human Daudi cell line - based assay or an antiviral activity of at least about  $2.1 \times 10^7$  units/milligram in a human WISH cell/EMCV-based assay.

**54.** The polypeptide of claim 31, 34, 37, 41, or 51, further comprising a secretion/localization sequence.

**10 55.** The polypeptide of claim 31, 34, 37, 41, or 51, further comprising a polypeptide purification subsequence.

**56.** The polypeptide of claim 55, wherein the sequence that facilitates purification is selected from the group consisting of: an epitope tag, a FLAG tag, a polyhistidine tag, and a GST fusion.

**15 57.** The polypeptide of claim 31, 34, 37, 41, or 51, further comprising a Met at the N-terminus.

**58.** The polypeptide of claim 31, 34, 37, 41, or 51, comprising a modified amino acid.

**20 59.** The polypeptide of claim 58, wherein the modified amino acid is selected from the group consisting of: a glycosylated amino acid, a PEGylated amino acid, a farnesylated amino acid, an acetylated amino acid, and a biotinylated amino acid.

**60.** A composition comprising the polypeptide of claim 31, 34, 37, 41, or 51 and an excipient.

**25 61.** The composition of claim 60, wherein the excipient is a pharmaceutically acceptable excipient.

**62.** A composition comprising the polypeptide of claim 58 in a pharmaceutically acceptable excipient.

**63.** A polypeptide which is specifically bound by a polyclonal antisera raised against at least one antigen, said at least one antigen comprising at least one amino acid sequence of SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, or a fragment thereof, wherein the antisera is subtracted with an IFN-alpha polypeptide encoded by a nucleic acid corresponding to one or more of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).

**64.** An antibody or antisera produced by administering the polypeptide of claim 31, 34, 37, 41, or 51 to a mammal, which antibody or antisera specifically binds at least one antigen, said at least one antigen comprising a polypeptide comprising one or more of the amino acid sequences of SEQ ID NO:36 to SEQ ID NO:70 and SEQ ID NO:79 to SEQ ID NO:85, or a fragment thereof, which antibody or antisera does not specifically bind to an IFN- $\alpha$  polypeptide encoded by a nucleic acid corresponding to one or more of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).

**65.** An antibody or antisera which specifically binds a polypeptide, the polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, wherein the antibody or antisera does not specifically bind to an IFN-alpha polypeptide encoded by a nucleic acid corresponding to one or more of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14),

V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).

5           **66.** A method of producing a polypeptide, the method comprising:  
introducing into a population of cells a nucleic acid of claim 1, 2, 3, 8, or 18, the  
nucleic acid operatively linked to a regulatory sequence effective to produce the encoded  
polypeptide; and  
culturing the cells in a culture medium to produce the polypeptide.

10          **67.** A method of producing a polypeptide, the method comprising:  
introducing into a population of cells a recombinant expression vector comprising  
the nucleic acid of claim 1, 2, 3, 8, or 18; and  
culturing the cells in a culture medium under conditions suitable to produce the  
polypeptide encoded by the expression vector.

15          **68.** A method of inhibiting growth of population of tumor cells, the  
method comprising:  
contacting the population of tumor cells with an effective amount of a polypeptide  
of claim 31, 34, 37, 41, or 51 sufficient to inhibit growth of tumor cells in said population  
of tumor cells, thereby inhibiting growth of tumor cells in said population of cells.

20          **69.** The method of claim 68, wherein the tumor cells are selected from the  
group consisting of: human carcinoma cells, human leukemia cells, human T-lymphoma  
cells, and human melanoma cells.

**70.** The method of claim 68, wherein the tumor cells are in culture.

25          **71.** A method of inhibiting the replication of a virus within at least one  
cell infected by the virus, the method comprising:  
contacting said at least one infected cell with an effective amount of a polypeptide  
of claim 31, 34, 37, 41, or 51 sufficient to inhibit viral replication in said at least one  
infected cell, thereby inhibiting replication of the virus in said at least one infected cells.

72. The method of claim 71, wherein the virus is an RNA virus.
73. The method of claim 72, wherein the virus is a human immunodeficiency virus or a hepatitis C virus.
74. The method of claim 71, wherein the virus is a DNA virus.
- 5 75. The method of claim 74, wherein the virus is a hepatitis B virus.
76. The method of claim 71, wherein the cells are cultured.
77. A method of treating an autoimmune disorder in a patient, the method comprising: administering to the patient an effective amount of the polypeptide of claim 31, 34, 37, 41, or 51.
- 10 78. The method of claim 77, wherein the autoimmune disorder is selected from the group consisting of multiple sclerosis, rheumatoid arthritis, lupus erythematosus, and type I diabetes.
- 15 79. In a method of treating a disorder treatable by administration of interferon-alpha to a subject, an improved method comprising: administering to the subject an effective amount of the polypeptide of claim 31, 34, 37, 41, or 51.
80. The method of claim 79, wherein the disorder treatable by administration of interferon-alpha is selected from the group consisting of: sclerosis, rheumatoid arthritis, lupus erythematosus, and type I diabetes.
- 20 81. A method of for making a modified or recombinant nucleic acid, the method comprising:  
recursively recombining a sequence of one or more nucleic acids of claim 1, 2, 3, 8, or 18 with a sequence of one or more additional nucleic acids, each sequence of the one or more additional nucleic acids encoding an interferon-alpha or an amino acid subsequence thereof.
- 25 82. The method of claim 81, wherein said recursive recombination produces at least one library of recombinant interferon-alpha homologue nucleic acids.

83. A nucleic acid library produced by the method of claim 82.
84. A population of cells comprising the library of claim 83.
85. A recombinant interferon-alpha homologue nucleic acid produced by the method of claim 82.
- 5        86. A cell comprising the nucleic acid of claim 85.
87. The method of claim 81, wherein the recursive recombination is performed *in vitro*.
88. The method of claim 81, wherein the recursive recombination is performed *in vivo* or *ex vivo*.
- 10        89. A composition comprising two or more nucleic acids of claim 1, 2, 3, 8, or 18.
90. The composition of claim 89, wherein the composition comprises a library comprising at least ten nucleic acids.
- 15        91. A method of producing a modified or recombinant interferon-alpha homologue nucleic acid comprising mutating a nucleic acid of claim 1, 2, 3, 8, or 18.
92. The modified or recombinant interferon-alpha homologue nucleic acid produced by the method of claim 91.
- 20        93. A computer or computer readable medium comprising a database comprising a sequence record comprising one or more character strings corresponding to a nucleic acid or protein sequence selected from SEQ ID NO:1 to SEQ ID NO:85.
94. An integrated system comprising a computer or computer readable medium comprising a database comprising one or more sequence records, each of said sequence records comprising one or more character strings corresponding to a nucleic acid or protein sequence selected from SEQ ID NO:1 to SEQ ID NO:85, the integrated system further comprising a user input interface allowing a user to selectively view said one or more sequence records.

**95.** The integrated system of claim 94, the computer or computer readable medium comprising an alignment instruction set which aligns the character strings with one or more additional character strings corresponding to a nucleic acid or protein sequence.

5       **96.** The integrated system of claim 95, wherein the instruction set comprises one or more of: a local homology comparison determination, a homology alignment determination, a search for similarity determination, and a BLAST determination.

10      **97.** The integrated system of claim 95, further comprising a user readable output element which displays an alignment produced by the alignment instruction set.

**98.** The integrated system of claim 94, the computer or computer readable medium further comprising an instruction set which translates at least one nucleic acid sequence comprising a sequence selected from SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78 into an amino acid sequence.

15      **99.** The integrated system of claim 94, the computer or computer readable medium further comprising an instruction set for reverse-translating at least one amino acid sequence comprising a sequence selected from SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85 into a nucleic acid sequence.

20      **100.** The integrated system of claim 99, wherein the instruction set selects the nucleic acid sequence by applying a codon usage instruction set or an instruction set which determines sequence identity to a test nucleic acid sequence.

25      **101.** A method of using a computer system to present information pertaining to at least one of a plurality of sequence records stored in a database, said sequence records each comprising at least one character string corresponding to SEQ ID NO:1 to SEQ ID NO:85, the method comprising:

          determining a list of at least one character string corresponding to one or more of SEQ ID NO:1 to SEQ ID NO:85 or a subsequence thereof;

          determining which of said at least one character string of said list are selected by a user; and

displaying each of the selected character strings, or aligning each of the selected character strings with an additional character string.

**102. The method of claim 101, further comprising displaying an alignment of each of the selected character strings with the additional character string.**

5           **103. The method of claim 101, further comprising displaying the list.**

**104. A nucleic acid which comprises a unique subsequence in a nucleic acid selected from SEQ ID NO:1 to SEQ ID NO:35 or SEQ ID NO:72 to SEQ ID NO:78, wherein the unique subsequence is unique as compared to a nucleic acid sequence of a known interferon-alpha nucleic acid sequence or a nucleic acid corresponding to any of**

10 **GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).**

15           **105. A polypeptide which comprises a unique subsequence in a polypeptide selected from: SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, wherein the unique subsequence is unique as compared to a sequence of a known interferon-alpha polypeptide or a sequence of a polypeptide encoded by a nucleic acid corresponding to any of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).**

25           **106. A target nucleic acid which hybridizes under stringent conditions to a unique coding oligonucleotide which encodes a unique subsequence in a polypeptide selected from: SEQ ID NO:36 to SEQ ID NO:70 or SEQ ID NO:79 to SEQ ID NO:85, wherein the unique subsequence is unique as compared to a sequence of a known interferon-alpha polypeptide or a sequence of a polypeptide encoded by a nucleic acid**

corresponding to any of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1).

107. The nucleic acid of claim 106, wherein the stringent conditions are selected such that a perfectly complementary oligonucleotide to the unique coding oligonucleotide hybridizes to the unique coding oligonucleotide with at least a 5x higher signal to noise ratio than for hybridization of the perfectly complementary oligonucleotide to a control nucleic acid corresponding to any of GenBank accession number: J00210 (alpha-D), J00207 (Alpha-A), X02958 (Alpha-6), X02956 (Alpha-5), V00533 (alpha-H), V00542 (alpha-14), V00545 (IFN-1B), X03125 (alpha-8), X02957 (alpha-16), V00540 (alpha-21), X02955 (alpha-4b), V00532 (alpha-C), X02960 (alpha-7), X02961 (alpha-10 pseudogene), R0067 (Gx-1), I01614, I01787, I07821, M12350 (alpha-F), M38289, V00549 (alpha-2a), and I08313 (alpha-Con1), wherein the target nucleic acid hybridizes to the unique coding oligonucleotide with at least a 2x higher signal to noise ratio as compared to hybridization of the control nucleic acid to the coding oligonucleotide.

108. The nucleic acid of any of claims 1, 2, 3, 8, or 18, wherein the nucleic acid encodes an interferon-alpha homologue having an increased growth inhibition activity against a population of cancer cells relative to a growth inhibition activity of human interferon-alpha 2a against said population of cancer cells.

109. The nucleic acid of claim 108, wherein the cancer cells of said population of cancer cells comprise a cancer cell line selected from: a leukemia cell line, a melanoma cell line, a lung cancer cell line, a colon cancer cell line, a central nervous system (CNS) cancer cell line, an ovarian cancer cell line, a breast cancer cell line, a prostate cancer cell line, and a renal cancer cell line, and the growth inhibition activity is measured as a concentration of interferon-alpha homologue producing a 50% inhibition of growth of the cancer cell line (GI50 value), wherein the interferon-alpha homologue has a GI50 value at least 2-fold lower than the GI50 value of the human interferon-alpha 2a.

**110.** The nucleic acid of claim 109, wherein the encoded interferon-alpha homologue has a GI50 value at least 5-fold lower than the GI50 value of the human interferon-alpha 2a.

5       **111.** The nucleic acid of claim 107, wherein the encoded interferon-alpha homologue has a GI50 value at least 10-fold lower than the GI50 value of the human interferon-alpha 2a.

10      **112.** The nucleic acid of any of claims 1, 2, 3, 8, or 18, wherein the nucleic acid encodes an interferon-alpha homologue having increased an cytostatic activity against a population of cancer cells relative to the cytostatic activity of human interferon-alpha 2a against said population of cancer cells.

15      **113.** The nucleic acid of claim 112, wherein the cancer cells comprise a cancer cell line selected from: a leukemia cell line, a melanoma cell line, a lung cancer cell line, a colon cancer cell line, a CNS cancer cell line, an ovarian cancer cell line, a breast cancer cell line, a prostate cancer cell line, and a renal cancer cell line, the cytostatic activity measured as the concentration of an interferon-alpha causing a total inhibition of growth of the cell line (TGI value), wherein the interferon-alpha homologue has a TGI value at least 2-fold lower than the TGI value of the human interferon-alpha 2a.

20      **114.** The nucleic acid of claim 112, wherein the encoded interferon-alpha homologue has a TGI value at least 5-fold lower than the TGI value of the human interferon-alpha 2a.

**115.** The nucleic acid of claim 112, wherein the encoded interferon-alpha homologue has a TGI value at least 10-fold lower than the TGI value of the human interferon-alpha 2a.

25      **116.** The nucleic acid of any of claims 1, 2, 3, 8, or 18, wherein the nucleic acid encodes an interferon-alpha homologue having an increased cytotoxic activity against a population of cancer cells relative to the cytotoxic activity of human interferon-alpha 2a against said population of cancer cells.

117. The nucleic acid of claim 116, wherein the cancer cells comprise a cancer cell line selected from: a leukemia cell line, a melanoma cell line, a lung cancer cell line, a colon cancer cell line, a central nervous system (CNS) cancer cell line, an ovarian cancer cell line, a breast cancer cell line, a prostate cancer cell line, and a renal cancer cell line, the cytotoxic activity measured as the concentration of interferon-alpha producing a 50% reduction in an amount of cellular protein in a cell line measured after a period of incubation (LC50 value), wherein the interferon-alpha homologue has a LC50 value at least 2-fold lower than the LC50 value of the human interferon-alpha 2a.

118. The nucleic acid of claim 116, wherein the encoded interferon-alpha homologue has a LC50 value at least 5-fold lower than the LC50 value of the human interferon-alpha 2a.

119. The nucleic acid of claim 116, wherein the encoded interferon-alpha homologue has a LC50 value at least 10-fold lower than the LC50 value of the human interferon-alpha 2a.

120. The polypeptide of any of claims claim 31, 34, 37, 41, or 51, said polypeptide having an increased growth inhibition activity against a population of cancer cells relative to the inhibition activity of human interferon-alpha 2a against the population of cancer cells.

121. The polypeptide of claim 120, wherein the population of cancer cells comprises a cancer cell line selected from: a leukemia cell line, a melanoma cell line, a lung cancer cell line, a colon cancer cell line, a CNS cancer cell line, an ovarian cancer cell line, a breast cancer cell line, a prostate cancer cell line, and a renal cancer cell line, the growth inhibition activity measured as the concentration of polypeptide or human interferon-alpha 2a causing a 50% inhibition of growth of the cell line (GI50 value), wherein the polypeptide has a GI50 value at least 2-fold lower than the GI50 value of the human interferon-alpha 2a.

122. A nucleic acid produced by the method of claim 81.

123. An interferon-alpha polypeptide or amino acid subsequence thereof produced by the method of claim 81.

1 40

SEQ\_36 (1) CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGFPQ  
SEQ\_37 (1) CDLPQTHSLGDRRAMILLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_38 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_39 (1) CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGFPQ  
SEQ\_40 (1) CDLPQTHSLGNRRALVLLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_41 (1) CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_42 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRQDFGFPQ  
SEQ\_43 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_44 (1) CDLPQTHSLGNRRALILLAQMRRISPFSCLKDRHDFGFPQ  
SEQ\_45 (1) CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRQDFGFPQ  
SEQ\_46 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_47 (1) CDLPQTHSLGNRRALILLGQMGRISHFSCLKDRHDFGFPQ  
SEQ\_48 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_49 (1) CDLPQTHSLGNRRALMLLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_50 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGLPQ  
SEQ\_51 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_52 (1) CDLPQTHSLGNKRAMMLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_53 (1) CDLPQTHSLGNSRALMLLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_54 (1) CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRHDFGFPQ  
SEQ\_55 (1) CDLPQTHSLGNRRAMMLAQMRSRISPSSCLMDRHDFFEPQ  
SEQ\_56 (1) CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRYDFGFPQ  
SEQ\_57 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFRFPQ  
SEQ\_58 (1) CDLPQTHSLGNRRRTLMIMAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_59 (1) CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRYDFGFPQ  
SEQ\_60 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_61 (1) CDLPQTHSLGNRRALILLAQMRRISPFSCLKDRHDFGFPQ  
SEQ\_62 (1) CDLPQTHSLGNRRALILLAQMGRVSPFSCLKDRHDFGFPQ  
SEQ\_63 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFRFPQ  
SEQ\_64 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_65 (1) CDLPQTHSLGNRRPLILLAQMGRISPFSCLKDRQDFGFPQ  
SEQ\_66 (1) CDLPQTHSPGNRRALMLLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_67 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGLPQ  
SEQ\_68 (1) CDLPQTHSLGNRRRTLMLMAQMRRISPFPRLKDRYDFGFPQ  
SEQ\_69 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_70 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRYDFGFPQ  
SEQ\_79 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLMDRHDFFGPQ  
SEQ\_80 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_81 (1) CDLPQTHSLGNRRRTLIMAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_82 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_83 (1) CDLPQTHSLGNRRRTLIMAQMGRISPFSCLKDRHDFGFPQ  
SEQ\_84 (1) CDLPQTHSLGNRRALILLAQMGRISPFSCLMDRHDFFGPQ  
SEQ\_85 (1) CDLPQTHSLGNRRRTLIMAQMGRISPFSCLKDRHDFGFPQ

**Fig. 1A**

41

80

SEQ_36	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQT
SEQ_37	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_38	(41) EEFDGNQFQKAQAI SVLHEMMQQTFNLFSTKNSSAAWDET
SEQ_39	(41) EEFDSNQFQKAQAI SVLHEMMQQTFNLFSTKDSSAAWDET
SEQ_40	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWDET
SEQ_41	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWDET
SEQ_42	(41) EEFDGNRFQKAQAI SVLHEMIQQTFNLFSTKNSSAAWEQS
SEQ_43	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_44	(41) EEFDSNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_45	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_46	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_47	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSVAWDER
SEQ_48	(41) EEFDGNQFQKAQAI SVLHEIMQQTFNLFSTKNSSAAWDET
SEQ_49	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_50	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKNSSAAWDET
SEQ_51	(41) EEFDGNQFQKAQAI SVLHEMMQQTFNLFSTKNSSAAWDET
SEQ_52	(41) EEFDGNQFQRAQAI FVLHEMIQQTFNFFSTKDSSAAWEQS
SEQ_53	(41) EEFDGNQFQKAQAI SAFHEMIQQTFNLFSTKDSSAAWEQN
SEQ_54	(41) EEFDGHQFQKTQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_55	(41) EEFDDKQFQKAPAI SVLHEVIQQTFNLFSTEDSSAAWEQT
SEQ_56	(41) EVFDGNQFQKAQAI SAFHEMMQQTFNLFSTEDSSAAWEQS
SEQ_57	(41) EEFDGNQLQKTQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_58	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_59	(41) EVFDGNQFQKAQAI SAFHEMIQQTFNLFSTKDSSATWEQS
SEQ_60	(41) EEFDGNQSQKAQAI SVLHEMIQQTFNLFSTKDSSDTWDAT
SEQ_61	(41) EEFDGNQFQKAQAI SAFHEMIQQTFNLFSTKDSSAAWEQS
SEQ_62	(41) EEFDGNQFQKAQAI SAFHEMIQQTFNLFSTKDSSATWEQS
SEQ_63	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_64	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_65	(41) EEFDGNQFQKAQAI SVLHEMMQQTFNLFSTKNSSAAWEQS
SEQ_66	(41) GEFDGNQFQKAQAI SVLHEMMQQTFNLFSTKDSSAAWEQS
SEQ_67	(41) EEFDGNQFQKTQAI SVLHEMIQQTFNLFSTKDSSDTWEQS
SEQ_68	(41) EVFDGNQFQKAQAI FLFHEMMQQTFNLFSTKNSSAAWDET
SEQ_69	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWEQS
SEQ_70	(41) EEFDGNQLQKAQAI SVLHEMIQQTFNLFSTKDSSAAWEQS
SEQ_79	(41) EEFDDNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWDET
SEQ_80	(41) EEFDGNQFQKAQGI SVLHEMIQQTFHLFSTKDSSATWEQS
SEQ_81	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWDET
SEQ_82	(41) EEFGGNQFQKAQAI SVLHEMIQQTFNLFSTEDSSAAWDET
SEQ_83	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWDET
SEQ_84	(41) EEFDDNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWDET
SEQ_85	(41) EEFDGNQFQKAQAI SVLHEMIQQTFNLFSTKDSSATWDET

**Fig. 1B****2/14**

81

120

SEQ_36	(81) LLEKFSTELYQQLNNDLEACVIQEVGVKETPLMNVDISILAV
SEQ_37	(81) LLEKFSTELYQQLNELEACVIQEVGVGGETPLMNGDSILAV
SEQ_38	(81) LLEKFSTELYQQLNELEACVIQEVGVEETPLMNEDSILAV
SEQ_39	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_40	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_41	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_42	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_43	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLMNVDPIAV
SEQ_44	(81) LLEKFSTELHQQLNELEACVVQEVGVEETPLMNEDSILAV
SEQ_45	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_46	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_47	(81) LLDKLYTELYQQLNNDLEACVMQEWWGGTPLMNEDSILAV
SEQ_48	(81) LLEKFSTELYQQLNELEACVIQGVGVEETPLMNEDSILAV
SEQ_49	(81) LLEKFSTGLYQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_50	(81) LLEKFSTELYQQLNNLEACVIQEVGMEETPLMNVDISILAV
SEQ_51	(81) LLEKFSTELYQQLNELEACVIQEVGVEETPLMNEDSILAV
SEQ_52	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_53	(81) LLEKFSTELYQQLNNLEACVIQEVGMEETPLMNVDISILAV
SEQ_54	(81) LLEKFSTELYQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_55	(81) LLEKFSTELYQQLNNDLEACVMQEERVGGETPLMNADSI
SEQ_56	(81) LLEKFSTELHQQLNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_57	(81) LLEKFSTELNQQLNNDLEACVIQGVGVEETPPMNVDISILAV
SEQ_58	(81) LLEKFSTELNQQLNNDLEACVIQEAGVEETPLMNVDISILAV
SEQ_59	(81) LLEKFSTELYQQLNNLEACVIQEVGVEETPLMNEDSILAV
SEQ_60	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_61	(81) LLEKFSTELYQQLNNLEACVIQEVGMEETPLMNEDSILAV
SEQ_62	(81) LLEKFSTELYQQLNNLEACVIQEVGVEETPLMNVDISILAV
SEQ_63	(81) LLEKFSTELYQQLNNLEACVIQEVGVEETPLMNVDISILAV
SEQ_64	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLVNVDISILAV
SEQ_65	(81) LLEKFSTELHQQLNELEACVIQEVGVEETPLMNVDISILAV
SEQ_66	(81) LLEKFSTELYRQLNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_67	(81) LLEKFYIELFQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_68	(81) LLDKFYTELYQQLNNDLEACVMQEGRVGETPLMNADSI
SEQ_69	(81) LLEKFSTELNQQLNNDLEACVTQEVGVEETPLMNEDSILAV
SEQ_70	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_79	(81) LLDKFYTELYQQLNNDLEACVIQEVGVEETPLMNEDSILAV
SEQ_80	(81) LLEKFSTELNQQLNNDLEACVIQEVGVEETPLMNVDISILAV
SEQ_81	(81) LLDKFYTELYQQLNNDLEACMMQEVGVEDTPLMNVDISILTV
SEQ_82	(81) LLDKFYIELFQQLNNDLEACVMQEERVGGETPLMNADSI
SEQ_83	(81) LLDKFYTELYQQLNNDLEACMIQEVGVEETPLMNEDSILAV
SEQ_84	(81) LLDKFYTELYQQLNNDLEACMMQEVGVEETPLMNVDISILTV
SEQ_85	(81) LLDKFYTELYQQLNNDLEACMMQEVGVEETPLMNEDSILAV

**Fig. 1C****3/14**

121

160

SEQ\_36 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_37 (121) KKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_38 (121) KKYFQRITLYLTEKKYSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_39 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_40 (121) KKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_41 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_42 (121) KKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_43 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_44 (121) KKYLQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_45 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_46 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_47 (121) RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_48 (121) RKYFQRITLYLTEKKYSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_49 (121) KKYFQRITLYLTEKKYSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_50 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_51 (121) KKYFQRITLYLTEKKYSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_52 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_53 (121) RKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_54 (121) KKYFQRITLYLMEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_55 (121) RKYFQRITLYLTKKYSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_56 (121) RKYFQRITLYLMEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_57 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_58 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_59 (121) RKYFQRITLYLMEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_60 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_61 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_62 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_63 (121) KKYFQRITLYLTERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_64 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_65 (121) KKYFQRITLYLIERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_66 (121) RKYFQRITLYLTEKKHSPCSWEVVRAEIMRSFSFSTNLQK  
SEQ\_67 (121) RKYFQRITLYLTEEKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_68 (121) KKYFRRITLYLTEKKYSPCAWEAVRAEIMRSFSFSTNLQK  
SEQ\_69 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_70 (121) KKYFQRITLYLTERKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_79 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_80 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_81 (121) RKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_82 (121) KKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_83 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_84 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK  
SEQ\_85 (121) KKYFRRITLYLTEKKYSPCAWEVVRAEIMRSFSFSTNLQK

**Fig. 1D**

161 166

SEQ\_36 (161) RLRRKE  
SEQ\_37 (161) RLRRKE  
SEQ\_38 (161) RLRRKE  
SEQ\_39 (161) RLRRKE  
SEQ\_40 (161) RLRRKE  
SEQ\_41 (161) RLRRKE  
SEQ\_42 (161) RLRRKE  
SEQ\_43 (161) RLRRKE  
SEQ\_44 (161) RLRRKE  
SEQ\_45 (161) RLRRKE  
SEQ\_46 (161) RLRRKE  
SEQ\_47 (161) RLRRKE  
SEQ\_48 (161) RLRRKE  
SEQ\_49 (161) RLRRKE  
SEQ\_50 (161) RLRRKE  
SEQ\_51 (161) RLRRKE  
SEQ\_52 (161) RLRRKE  
SEQ\_53 (161) RLRRKE  
SEQ\_54 (161) RLRRKE  
SEQ\_55 (161) RLRRKE  
SEQ\_56 (161) RLRRKE  
SEQ\_57 (161) RLRRKE  
SEQ\_58 (161) RLRRKE  
SEQ\_59 (161) RLRRKE  
SEQ\_60 (161) RLRRKE  
SEQ\_61 (161) RLRRKE  
SEQ\_62 (161) RLRRKE  
SEQ\_63 (161) RLRRKE  
SEQ\_64 (161) RLRRKE  
SEQ\_65 (161) RLRRKE  
SEQ\_66 (161) RLRRKE  
SEQ\_67 (161) RLRRKE  
SEQ\_68 (161) RLRRKE  
SEQ\_69 (161) RLRRKE  
SEQ\_70 (161) RLRRKE  
SEQ\_79 (161) RLRRKE  
SEQ\_80 (161) RLRRKE  
SEQ\_81 (161) RLRRKE  
SEQ\_82 (161) RLRRKE  
SEQ\_83 (161) RLRRKE  
SEQ\_84 (161) RLRRKE  
SEQ\_85 (161) RLRRKE

**Fig. 1E**

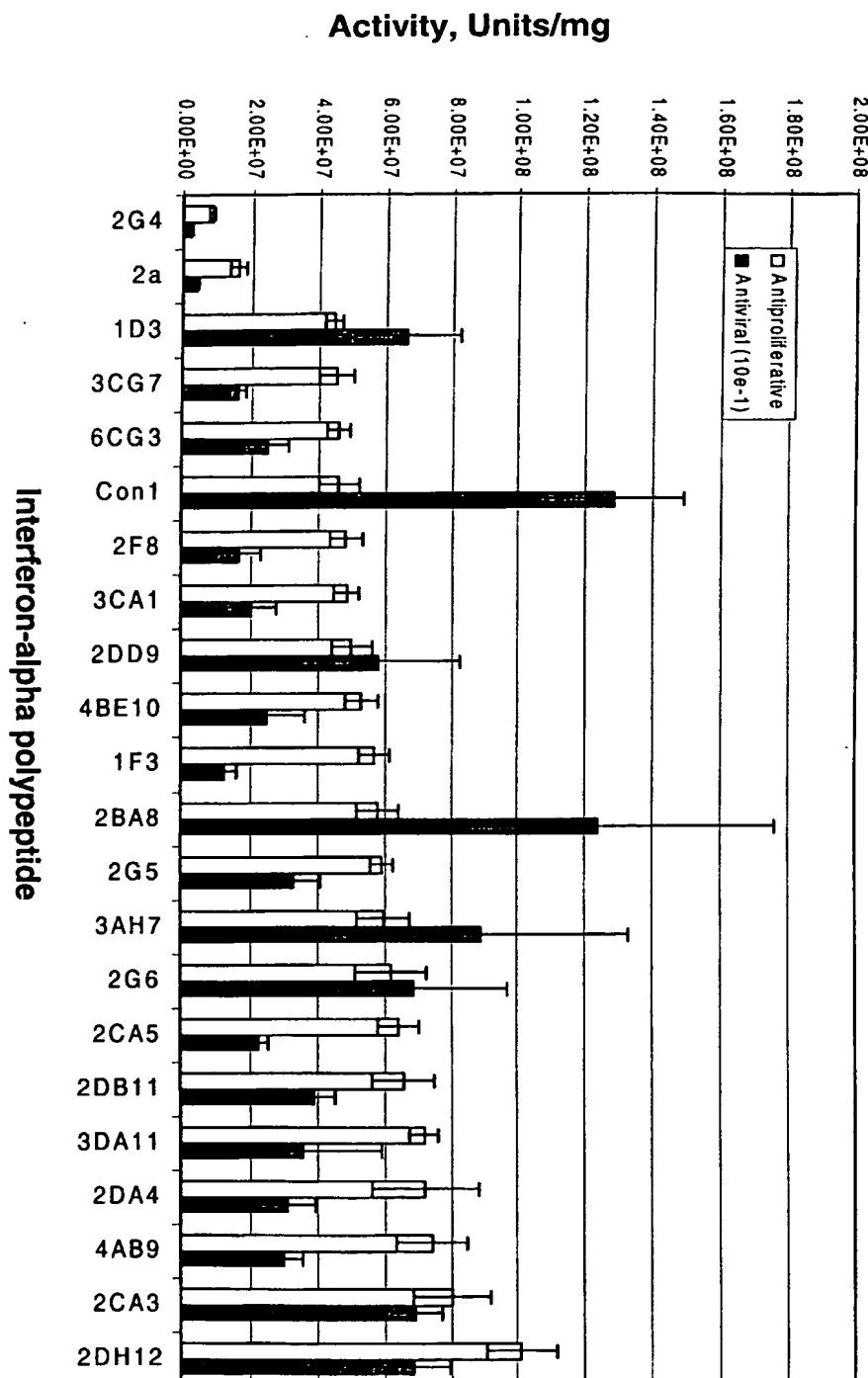


Fig. 2

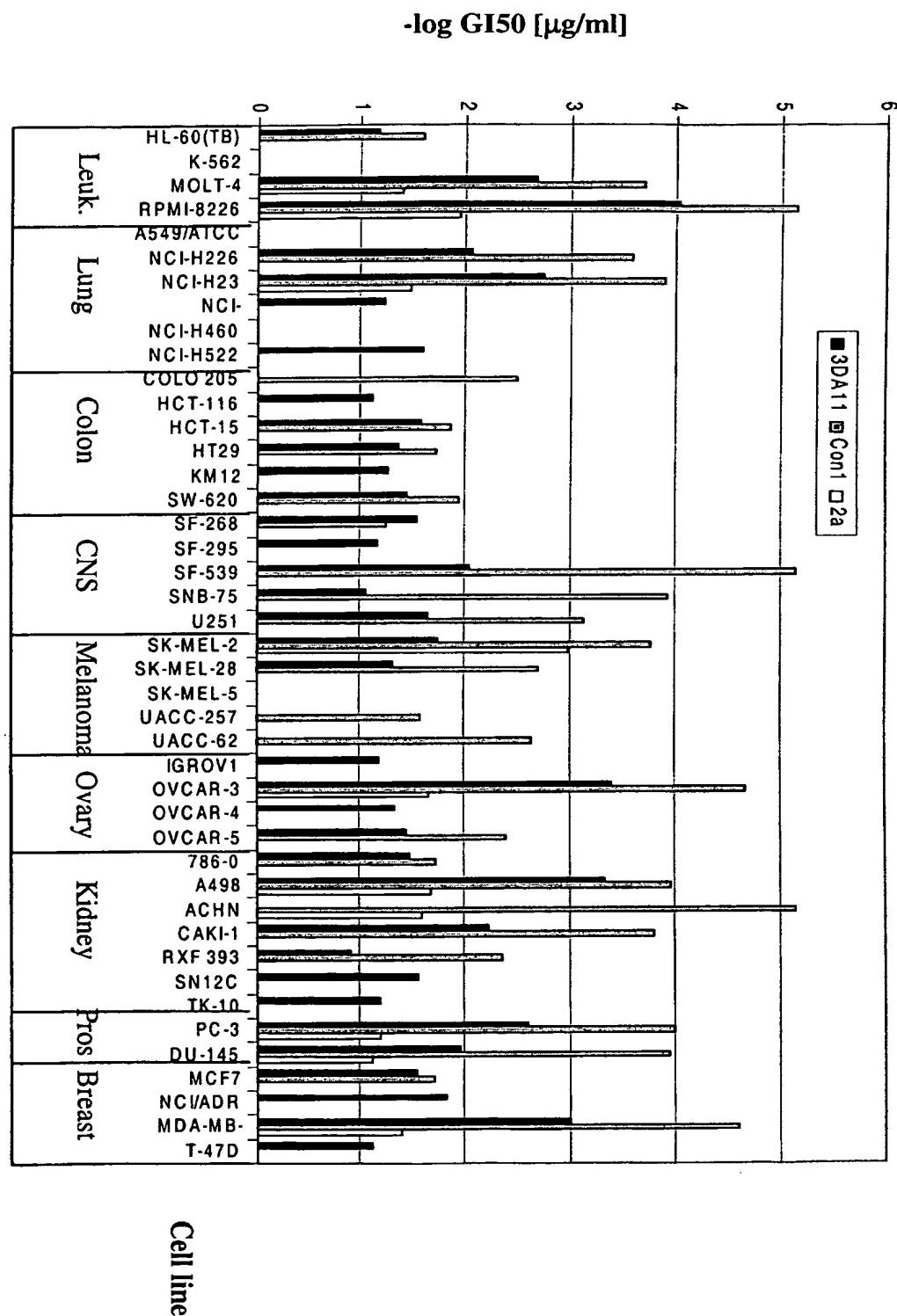


Fig. 3A

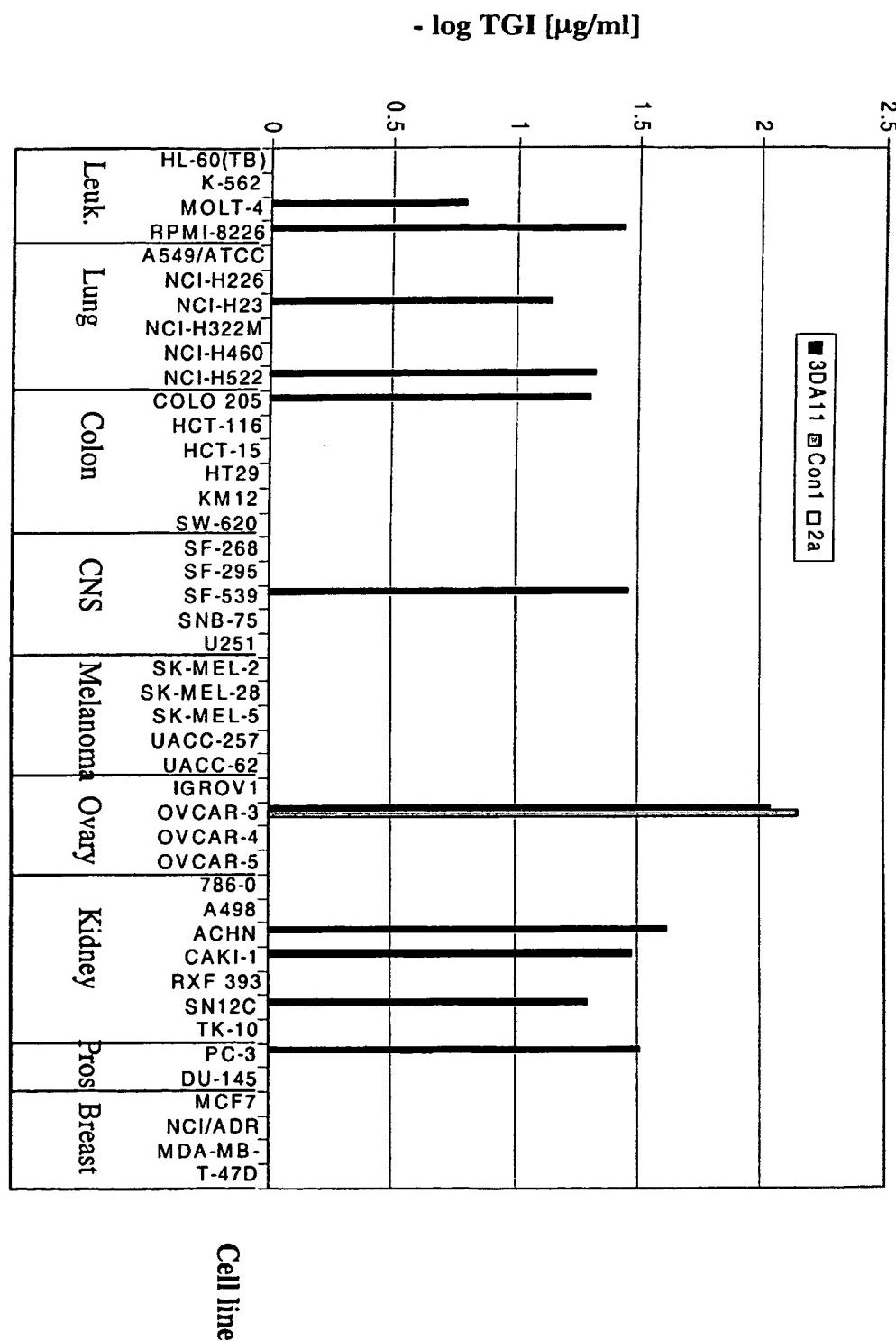


Fig. 3B

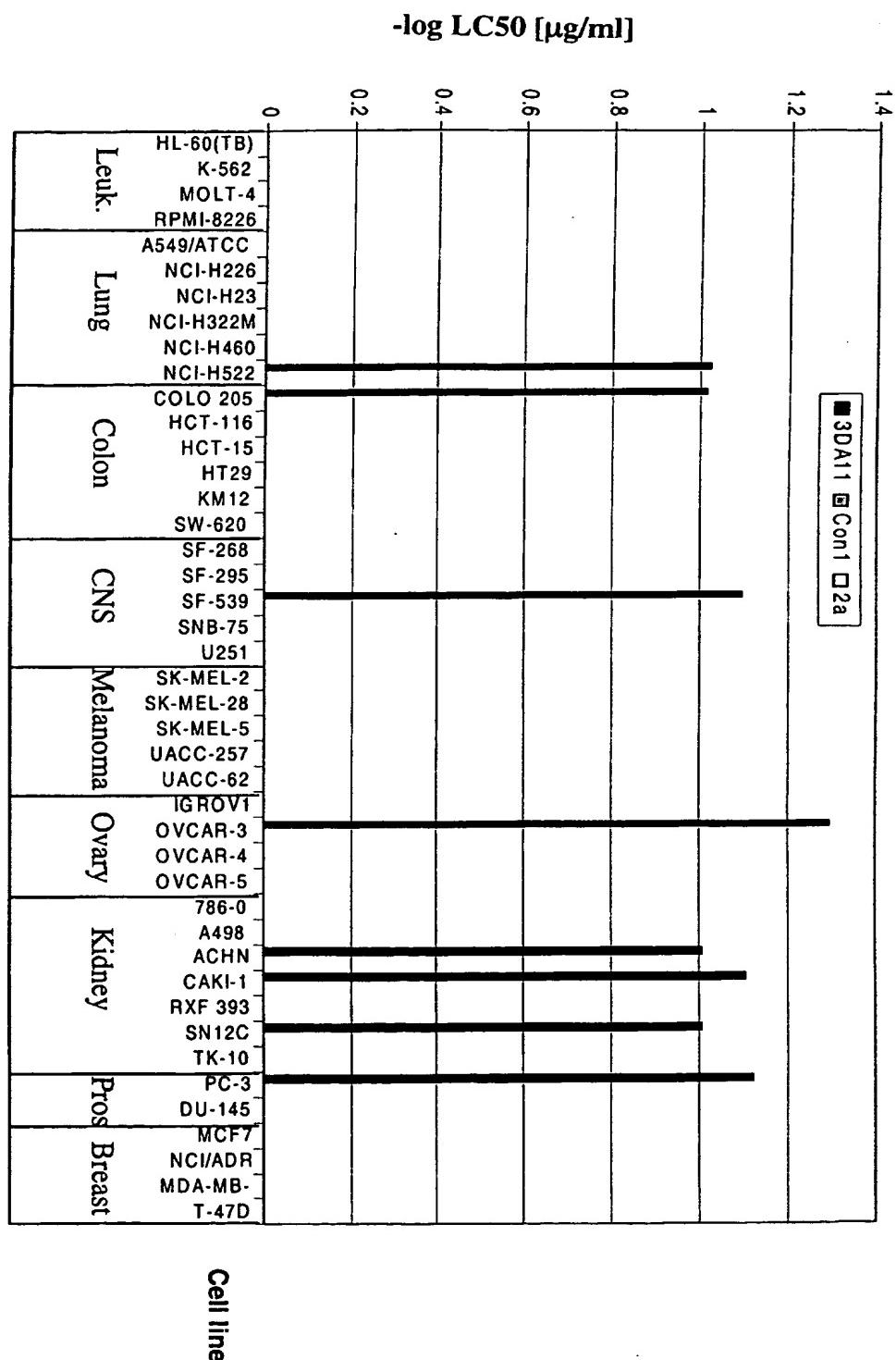


Fig. 3C

### Cytostatic activity on leukemia cell line (RPMI-8226)

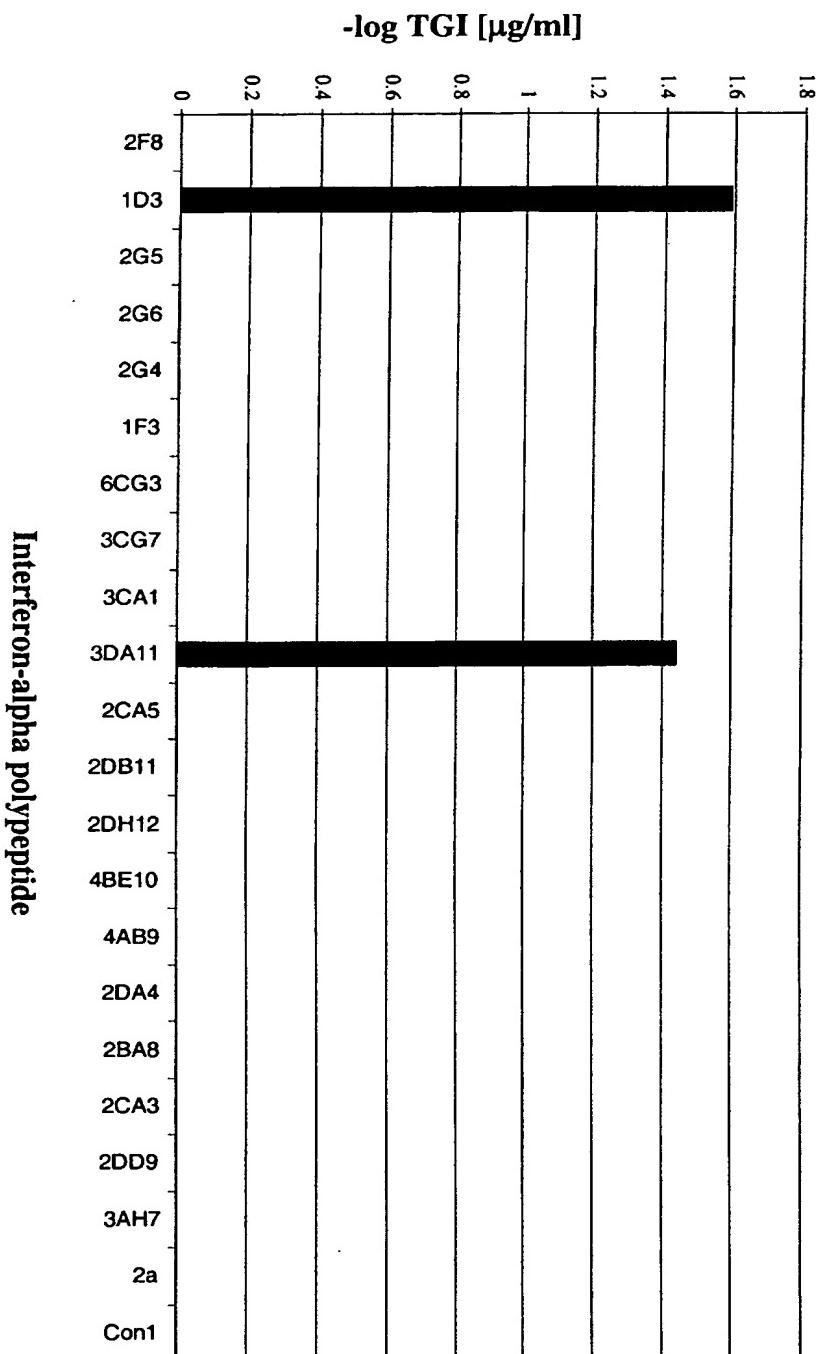


Fig. 4A

### Cytostatic activity on lung cancer cell line (NCI-H23)

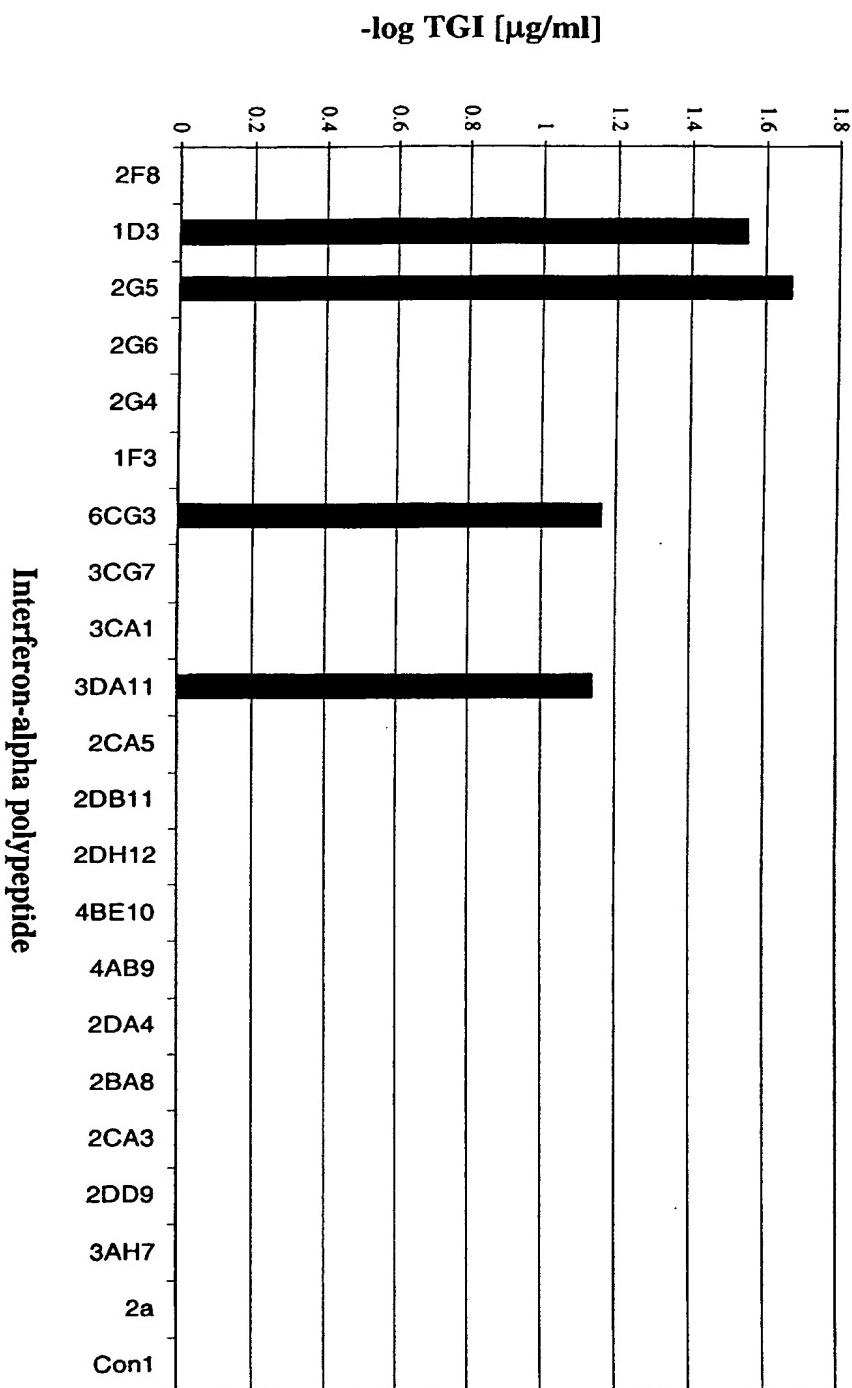
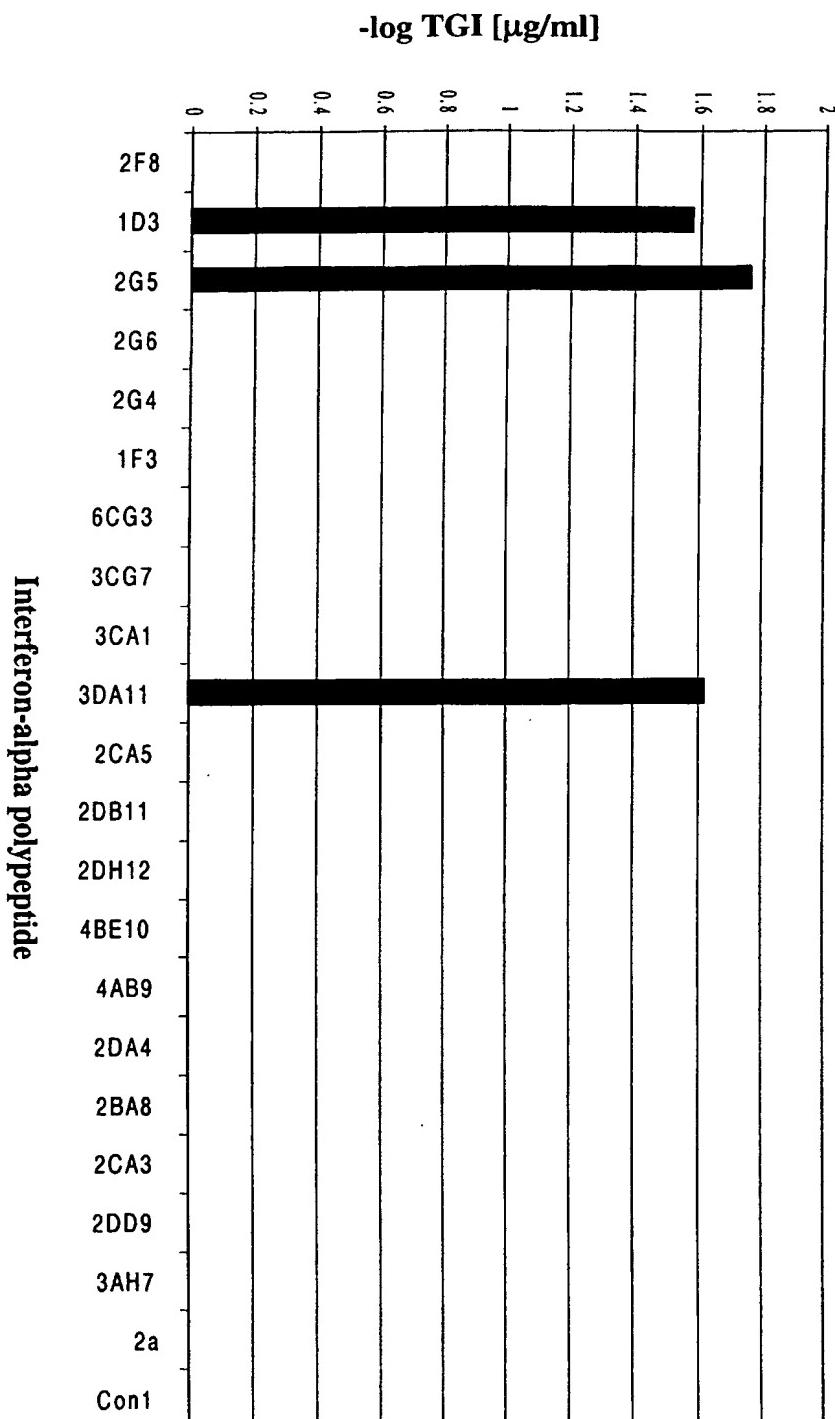


Fig. 4B

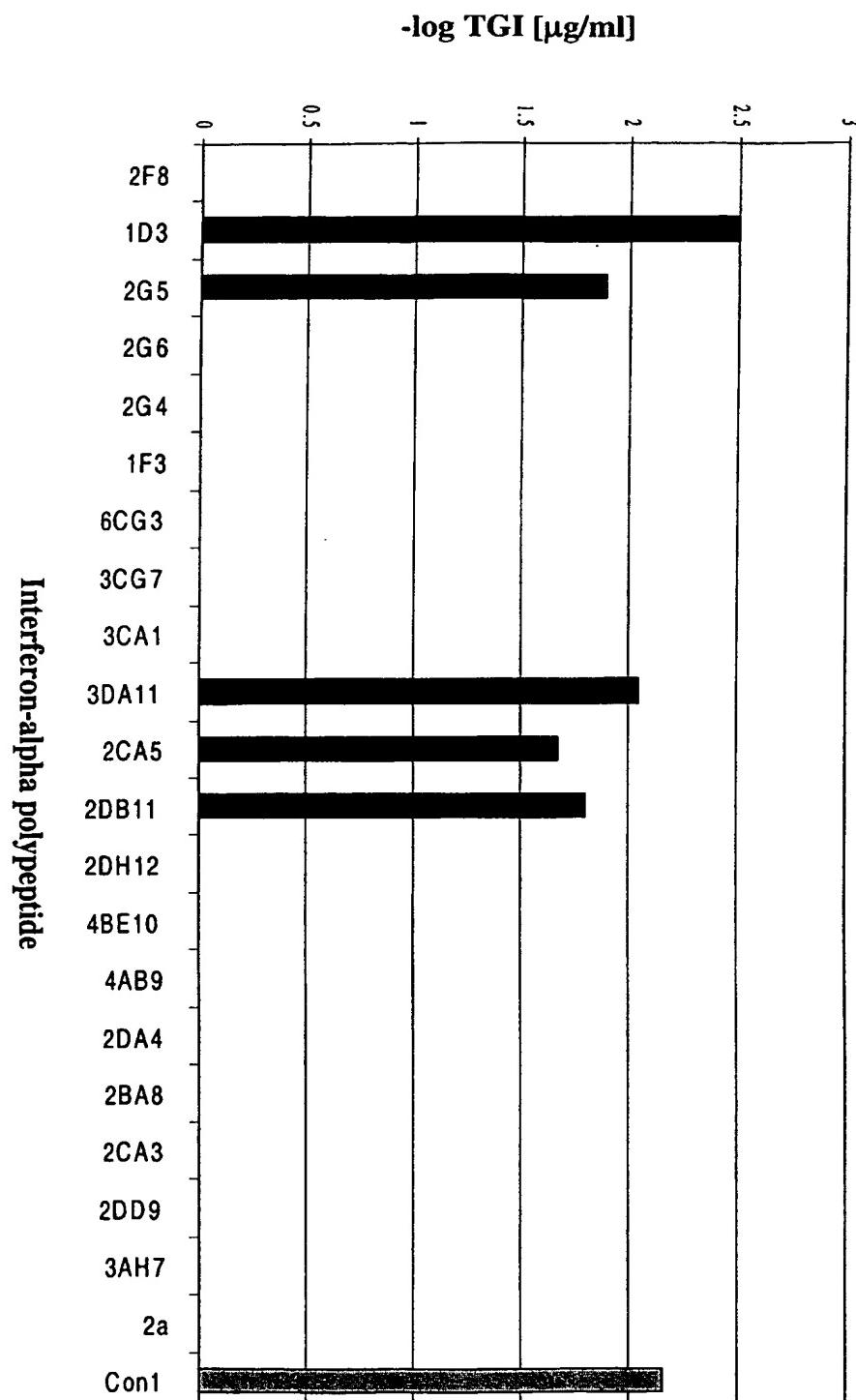
### Cytostatic activity on renal cancer cell line (ACHN)



**12/14**

Fig. 4C

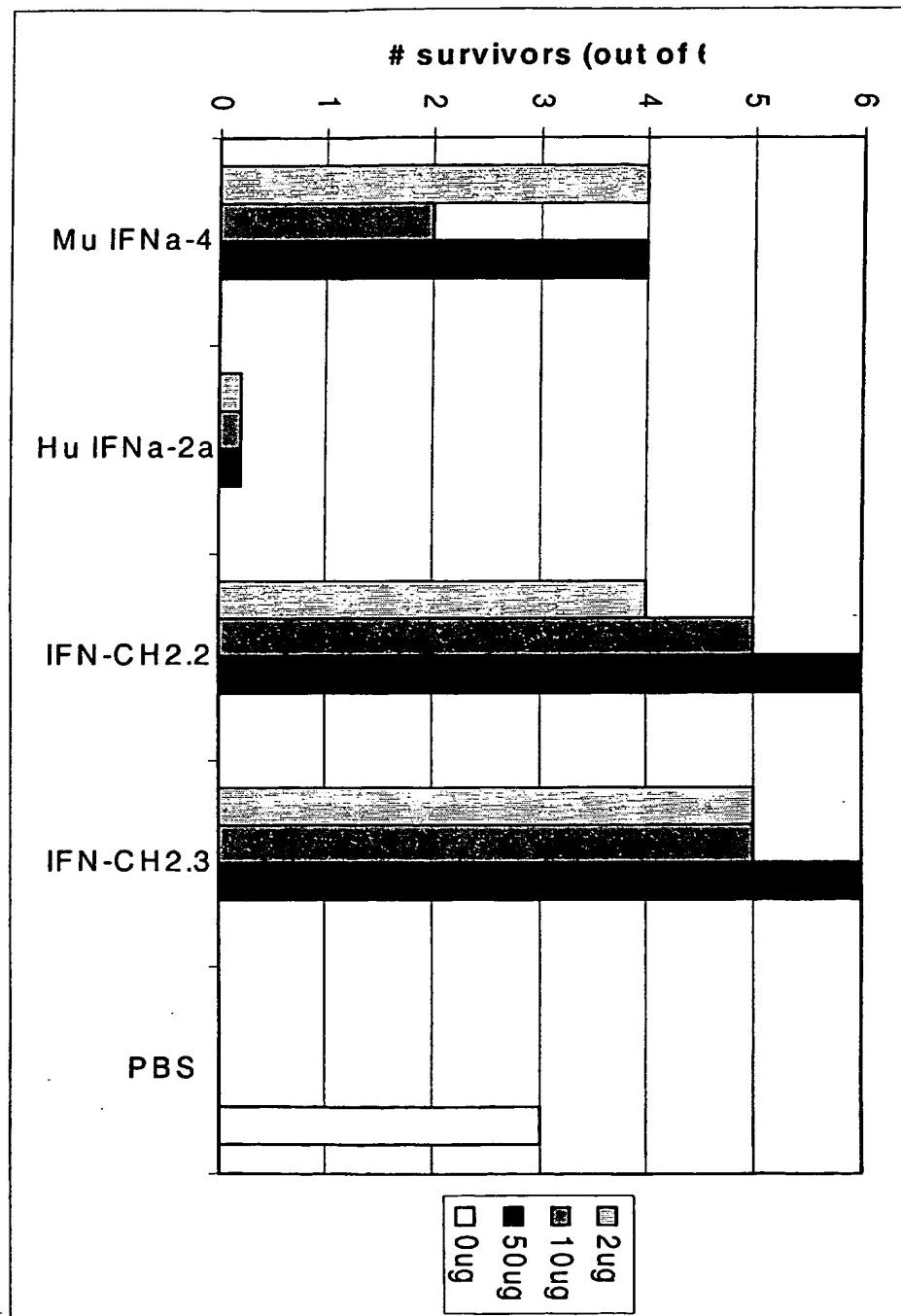
### Cytostatic activity on ovarian cancer cell line (OVCAR-3)



Interferon-alpha polypeptide

Fig. 4D

Fig. 5



14/14

**SEQUENCE LISTING**

5 <110> HEINRICHHS, VOLKER  
CHEN, TEDDY  
PATLEN, PHILLIP A.

10 <120> IFN-ALPHA HOMOLOGUES

15 <130> 02-101510/0140.002

<140>  
<141>

20 <150> 09/415,183  
<151> 1999-10-07

<160> 88

25 <170> PatentIn Ver. 2.0

<210> 1  
<211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

30 <220>  
<223> Clone ID 2DH12

<400> 1  
tgtgatctgc ctcagaccca cagccttggc aacaggaggg ctttgatgct cttggcacaaa 60  
atgggacgaa tctctccccc ctccctgcctg aaggacagac aagactttgg attccccccag 120  
gaggagttt atggcaacca gttccagaag gctcaaggca tctctgtcct ccatgagatg 180  
atccagcaga ctttcaatct cttcagcaca aaggattcat ctgctgttg ggaacagacc 240  
ctcttagaaa aattttccac tgaatctac cagcagctga atgacacttggaa agcctgcgtg 300  
atacaggagg taggggtgaa agagactccc ctgatgaatg tggactccat cttggctgtg 360  
40 aggaagtact tccaaagaat cacttttat ctaatagaga ggaaatacag cccttgcgtca 420  
tgggaggtt tcagagcaga aatcatgaga tctttctctt tttcaacaaa cttgcacaaa 480  
agattaagga ggaaggaa 498

<210> 2  
<211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

50 <220>  
<223> Clone ID 2CA3

55 <400> 2  
tgtgatctgc ctcagaccca cagccttggc gacaggaggg ccatgatact cttggcacaaa 60  
atgggacgaa tctctccccc ctccctgcctg aaggacagat atgatttcgg attccccccag 120  
qaggagttt atggcaacca gttccagaag gctcaaggca tctctgtcct ccatgagatg 180

atccagcaga cttcaatct cttcagcaca aaggattcat ctgctgcttg ggaacagagc 240  
ctcctagaaa aattttcac tgaacttac cagcagctga atgaactgga agcatgtgt 300  
atacaggagg ttggggtggg agagactccc ctgatgaatg gggactccat cctggctgtg 360  
aagaagtact tccaaagaat cacttttat ctaatagaga ggaaatacag cccttgc 420  
5 tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

<210> 3  
<211> 498  
10 <212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

15 <220>  
<223> Clone ID 4AB9

<400> 3  
20 tggatctgc ctcagaccca cagccttggc aacaggaggg cttgatact cctggcacaa 60  
atggacgaa tcttcctt ctcctgcctg aaggacagac atgactttgg attcccccg 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtctt ccatgagatg 180  
atgcagcaga cttcaatct cttcagcaca aagaactcat ctgctgcttg ggatgagacc 240  
25 ctctagaaa aattttcac tgaacttac cagcaactga atgaactgga agcatgtgt 300  
atacaggagg ttggggtggg agagactccc ctgatgaatg aggactccat cctggctgtg 360  
aagaataact tccaaagaat cacttttat ctgacagaga agaagtatag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

30 <210> 4  
<211> 498  
<212> DNA  
<213> Artificial Sequence

35 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2DA4

40 <400> 4  
tggatctgc ctcagaccca cagccttggt aacaggaggg cttgatgct cctggcacaa 60  
atggaaagaa tcttcctt ctcctgcctg aaggacagac aagactttgg attcccccg 120  
45 gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtctt ccatgagatg 180  
atgcagcaga cttcaatct cttcagcaca aaggactcat ctgctgcttg ggatgagacc 240  
ctctagaaa aattttcac tgaacttac cagcagctga atgacctgga agcctgcgtg 300  
atacaggagg ttggggtggg agagacccc ctgatgaatg tggactccat cctggctgtg 360  
50 agaagtact tccaaagaat cacttttat ctaatagaga ggaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

<210> 5  
<211> 498  
<212> DNA  
55 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 3DA11

5       <400> 5  
tgcgtatctgc ctcagaccca cagccttggt aacaggaggg ctttgtact cttggcacaa 60  
atggaaagaa tctctccccc ctcctgcctg aaggacat atgatttcgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtctt ccatgatg 180  
atccagcaga ctttcaatct cttcagcaca aaggattcat ctgctgctt ggatgagacc 240  
ctcttagaaa aatttccac tgaactttac cagcagctga atgacttggaa agcctgcgtg 300  
10      atacaggagg ttgggggttggaa agagactccc ctgatgaatg aggactccat cttggctgtg 360  
aagaataact tccaaagaat cacttttat ctaatagaga gaaatacag cccttgcgtca 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

15      <210> 6  
<211> 498  
<212> DNA  
<213> Artificial Sequence

20      <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2DB11

25      <400> 6  
tgcgtatctgc ctcagaccca cagccttggt aacaggaggg ctttgtact cttggcacaa 60  
atggaaagaa tctctccccc ctcctgcctg aaggacat atgatttcgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtctt ccatgatg 180  
30      atccagcaga ctttcaatct cttcagcaca aaggattcat ctgctgctt ggatgagacc 240  
ctcttagaaa aatttccac tgaactttac cagcagctga atgacttggaa agcctgcgtg 300  
atacaggagg ttgggggttggaa agagactccc ctgatgaatg tggactccat cttggctgtg 360  
aggaagtact tccaaagaat cacttttat ctaatagaga gaaatacag cccttgcgtca 420  
35      tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 7  
<211> 498  
<212> DNA  
40      <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

45      <220>  
<223> Clone ID 2CA5

<400> 7  
tgcgtatctgc ctcagaccca cagccttggt aacaggaggg ctttgtact cttggcacaa 60  
50      atggacgaa tctctccccc ctcctgcctg aaggacacac aagactttgg attccccca 120  
gaggagttt atggcaaccc gttccagaag gctcaagcca tctctgtctt ccatgatg 180  
atccagcaga ctttcaatct cttcagcaca aagaactcat ctgctgctt ggAACAGAGC 240  
ctcttagaaa aatttccac tgaactttac cagcagctga atgacttggaa agcctgcgtg 300  
atacaggagg ttgggggttggaa agagactccc ctgatgaatg aggactccat cttggctgtg 360  
55      agaaataact tccaaagaat cacttttat ctaatagaga gaaatacag cccttgcgtca 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 8

<211> 498  
<212> DNA  
<213> Artificial Sequence

5 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2G6

10 <400> 8  
tgtgatctgc ctcagaccca cagccttgggt aacaggaggg ctttgataact cctggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtccct ccatgagatg 180

15 atccagcaga ccttcaatct cttagcaca aaggactcat ctgctacttg ggaacagagc 240  
ctcctagaaa aattttccac tgaacttaac cagcagctga atgacctgga agcctgcgtg 300  
atacaggagg ttgggggttga agagactccc ctgatgaatg tggacccat cctggctgtg 360  
aagaataact tccaaagaat cactcttat ctgacagaga agaaatacag cccttgc 420  
tggaggtt gtcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
20 agattaagga ggaaggaa 498

<210> 9  
<211> 498  
<212> DNA  
25 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

30 <220>  
<223> Clone ID 3AH7

<400> 9  
tgtgatctgc ctcagaccca cagccttgggt aacaggaggg ctttgataact cctggcacaa 60  
35 atgcgaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtccct ccatgagatg 180  
atccagcaga ccttcaatct cttagcaca aaggattcat ctgctacttg ggaacagagc 240  
ctcctagaaa aattttccac tgaacttcac cagcaactga atgaactgga agcatgtgt 300  
gtacaggagg ttgggggttga agagactccc ctgatgaatg aggactccat cctggctgtg 360  
40 aagaataacc tccaaagaat cacttttat ctgacagaga agaagtatacg cccttgc 420  
tggaggtt gtcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 10  
45 <211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
50 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2G5

55 <400> 10  
tgtgatctgc ctcagaccca cagccttgggt aacaggaggg ctttgatgct cctggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac aagactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtccct ccatgagatg 180  
atccagcaga ccttcaatct cttagcaca aaggattcat ctgctacttg ggaacagagc 240

ctcctagaaa aattttccac tgaactctac cagcagctga atgacacctgga agcctgcgtg 300  
atacaggagg ttggggtgga agagaccccc ctgatgaatg tggactccat cctggctgtg 360  
aggaagtact tccaaagaat cactcttat ctaatagaga gaaatacag cccttgcgtca 420  
tgggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
5 agattaagga ggaaggaa 498

<210> 11  
<211> 498  
<212> DNA  
10 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

15 <220>  
<223> Clone ID 2BA8

<400> 11  
tgtgatctgc ctcagaccca cagccttggt aacaggaggg ccctgatact cctggcacaa 60  
20 atggaccaa tctctccccc ctcctgcctg aaggacagat atgatttcgg attcccccaag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ccttcaatct cttcagcaca aaggattcat ctgctgttg ggaacagagc 240  
ctcctagaaa aattttccac tgaactttac cagcagctga atgacacctgga agcctgcgtg 300  
atacaggagg ttggggtgga agagaccccc ctaatgaatg tggactccat cctggctgtg 360  
25 aggaagtact tccaaagaat cactcttat ctaatagaga gaaatacag cccttgcgtca 420  
tgggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 12  
30 <211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
35 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 1F3

40 <400> 12  
tgtgatctgc ctcagaccca cagccttggt aacaggaggg ccctgatact cctggcacaa 60  
atgggaagaa tctctcattt ctcctgcctg aaggacagac atgactttgg attcccccaag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ccttcaacct cttcagcaca aaggactcat ctggtgttg ggatgagagg 240  
45 cttctagaca aactctatac tgaactttac cagcagctga atgacacctgga agcctgtgtg 300  
atgcaggagg tttgggggg aggactccc ctgatgaatg aggactccat cctggctgtg 360  
agaaaataact tccaaagaat cactcttat ctgacagaga agaaatacag cccttgcgtca 420  
tgggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

50 <210> 13  
<211> 498  
<212> DNA  
<213> Artificial Sequence

55 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 4BE10

<400> 13

5 tggatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cttggcacag 60  
atggaccaa tctccctt ctcctgcctg aaggacagat atgatttcgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagata 180  
atgcagcaga ctttcaatct ctgcagcaca aagaactcat ctgctgttg ggatgagacc 240  
ctcctagaaa aattttccac tgaactttac cagcaactga atgaacttgg aacatgtgtg 300  
atacaggggg ttgggttgg aagactccc ctgatgaatg aggactccat cttggctgtg 360  
10 agaaaatact tccaaagaat cacttttat ctgacagaga agaagtatag cccttgttcc 420  
tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 14

15 <211> 498  
<212> DNA  
<213> Artificial Sequence

<220>

20 <223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 2DD9

25 <400> 14

tggatctgc ctcagaccca cagccttgg aacaggaggg ctttgatgt cttggcacaa 60  
atggaaagaa tctccctt ctcctgcctg aaggacagat atgatttcgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ctttcaatct ctgcagcaca aaggattcat ctgctgttg gaaacagagc 240  
30 ctcctagaaa aattttccac tggactctac cagcagctga atgacttgg aacctgcgtg 300  
atacaggagg ttgggttgg aagacccccc ctgatgaatg aggactccat cttggctgtg 360  
aagaaaatact tccaaagaat cacttttat ctgacagaga agaagtatag cccttgttcc 420  
tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

35

<210> 15

<211> 498  
<212> DNA  
<213> Artificial Sequence

40

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<220>

45 <223> Clone ID 3CA1

<400> 15

tggatctgc ctcagaccca cagccttggc aacaggaggg ctttgataact cttggcacaa 60  
atggaaagaa tctccctt ctcctgcctg aaggacagac atgactttgg attacccca 120  
50 gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ctttcaatct ctgcagcaca aagaactcat ctgctgttg ggatgagacc 240  
ctcctagaaa aattttccac tgaactttac cagcaactga ataaaccttgg aacatgtgtg 300  
atacaggagg ttgggttgg aagactccc ctgatgaatg tggactccat cttggctgtg 360  
aagaaaatact tccaaagaat cacttttat ctgacagaga agaagtatag cccttgtgcc 420  
55 tggaggttg tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 16

<211> 498

<212> DNA  
<213> Artificial Sequence

5 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2F8

10 <400> 16  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ccttgatact cctggcacaa 60  
atgggacgaa tctctccccc tccctgcctg aaggacagat atgatttcgg attccccccag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atgcagcaga ccttcaatct cttagcaca aagaactcat ctgctgctt ggtatgagacc 240  
15 ctcctagaaa aattttccac tgaactttac cagcaactga atgaactgga agcatgtgt 300  
atacaggagg ttgggggttga agagactccc ctgatgaatg aggactccat cctggctgt 360  
aagaataact tccaaagaat cacttttat ctgacagaga agaagatag cccttggcc 420  
tgggaggttgc ttagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

20 <210> 17  
<211> 498  
<212> DNA  
<213> Artificial Sequence

25 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

30 <223> Clone ID 6CG3

<400> 17  
tgtgatctgc ctcagaccca cagccttgg aacaagaggg ccatgatgct cctggcacaa 60  
atgggaaagaa cctctccccc tccctgcctg aaggacagac atgactttgg attccccccag 120  
35 gaggagttt atggcaacca gttccagagg gctcaagcca tctttgcctt ccatgagatg 180  
atccagcaga ccttcaatct cttagcaca aaggactcat ctgctgctt ggaacagagc 240  
ctcctagaaa aattttccac tgaacttaac cagcagctga atgacctgga agcctgcgt 300  
atacaggaag ttgggggttga agagactccc ctgatgaatg aggactccat cctggctgt 360  
aagaataact tccaaagaat cacttttat ctgacagaga agaaatacag cccttggcc 420  
40 tgggaggttgc ttagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 18  
<211> 498  
45 <212> DNA  
<213> Artificial Sequence

<220>

50 <223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 3CG7

<400> 18  
55 tgtgatctgc ctcagaccca cagccttgg aacagttaggg ctttgatgct cctggcacaa 60  
atgggaaagaa tctccccccc tccctgcctg aaggacagac atgatttcgg attccccccag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atccagcaga ccttcaatct cttagcaca aaggattcat ctgctgctt ggaacacagaac 240  
ctcctagaaa aattttccac tgaactttac cagcaactga ataacctgga agcatgtgt 300

atacaggagg ttgggatgga agagactccc ctgatgaatg tggactccat cctggctgtg 360  
aggaagtact tccaaagaat cacttttat ctaatagaga ggaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

5 <210> 19  
<211> 498  
<212> DNA  
<213> Artificial Sequence

10 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

15 15 <223> Clone ID 1D3

<400> 19  
tgtatctgc ctcagaccca cagccttgg aacaggaggg cttgtatact cctggcacaa 60  
atggaaatc tcttcattt ctccgtctg aaggacagac atgatttcgg attccccag 120  
20 gaggagttt atggccacca gttccagaag actcaagcca tctctgtctt ccatgagatg 180  
atccagcaga cttcaatct cttcagcaca aaggactcat ctgctgttg ggaacagagc 240  
ctccatggaaa aattttccac tgaactttac cagcaactga atgacctgga agcatgtgtg 300  
atacaggagg ttggggtgaa agagactccc ctgatgaatg aggactccat cctggctgtg 360  
25 aagaaatact tccaaagaat cacttttat ctgatggaga agaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

<210> 20  
<211> 498  
30 <212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

35 <223> Clone ID 2G4

<400> 20  
tgtatctgc ctcagaccca cagccttgg aacaggaggg ccatgtatgct cctggcacaa 60  
atgacgacaa tcttccttc ctccgtctg atggacagac atgacttga attccccag 120  
gaggatattt atgataaaaca gttccagaag gctccagcca tctctgtctt ccatgaggtg 180  
attcagcaga cttcaatct cttcagcaca aaggactcat ctgctgttg ggaacagacc 240  
45 ctccatggaaa aattttccac tgaactttac cagcaactga atgacctgga agcatgtgtg 300  
atgcaggagg agagggtggg agaaactccc ctgatgaatg cgactccat cttggctgtg 360  
agaaatact tccaaagaat cacttttat ctgacaaaga agaagtatag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgc 480  
agattaagga ggaaggaa 498

50 <210> 21  
<211> 498  
<212> DNA  
<213> Artificial Sequence

55 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 1A1

<400> 21  
5 tggatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cttggcacaa 60  
atgggaagaa tctctcattt ctccgcctg aaggacagat atgatttcgg attccccca 120  
gagggtgttg atggcaacca gttccagaag gcccaagcca tctctgcctt ccatgagatg 180  
atgcagcaga cttcaatct cttcagcaca gaggactcat ctgctgttg gaaacagagc 240  
ctcctagaaa aattttccac tgaacttcac cagcaactga atgacctgga agcctgttg 300  
atacaggagg ttgggggttga agagactccc ctgatgaatg aggactccat cctggctgtg 360  
10 agggaaatact ttcaaagaat cacttttat ctaatggaga agaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgcacaaaa 480  
agattaagga ggaaggaa 498

<210> 22  
<211> 498  
15 <212> DNA  
<213> Artificial Sequence

<220>  
20 <223> Description of Artificial Sequence: Synthetic DNA  
<220>  
<223> Clone ID 1D10

<400> 22  
25 tggatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cttggcacaa 60  
atgggaagaa tctctcattt ctccgcctg aaggacagac atgatttcgg attccccca 120  
gaggagttt atggccacca gttccagaag actcaagcca tctctgcctt ccatgagatg 180  
atccagcaga cttcaatct cttcagcaca aaggactcat ctgctgttg gaaacagagc 240  
ctcctagaaa aattttccac tgaacttac cagcaactga atgacctgga agcatgttg 300  
30 atacaggagg ttgggggttga agagactccc ctgatgaatg aggactccat cctggctgtg 360  
aaggaaatact ttcaaagaat cacttttat ctaatggaga agaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgcacaaaa 480  
agattaagga ggaaggaa 498

<210> 23  
<211> 498  
<212> DNA  
35 <213> Artificial Sequence

<220>  
40 <223> Description of Artificial Sequence: Synthetic DNA  
<220>  
<223> Clone ID 1F6

<400> 23  
45 tggatctgc ctcagaccca cagccttgg aacaggagga ctttgatgat aatggcacaa 60  
atgggaagaa tctctcattt ctccgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
50 atccagcaga cttcaatct cttcagcaca aaggactcat ctgctacttg gaaacagagc 240  
ctcctagaaa aattttccac tgaacttaac cagcagctga atgacctgga agcctgcgtg 300  
atacaggagg ctgggggttga agagactccc ctgatgaatg tggactccat cctggctgtg 360  
aaggaaatact ttcaaagaat cacttttat ctaacagaga agaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tcttctt tttcaacaaa ctgcacaaaa 480  
55 agattaagga ggaaggaa 498

<210> 24  
<211> 498  
<212> DNA

<213> Artificial Sequence

<220>

5    <223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 2A10

<400> 24

10   tttttttttt ctcagaccca cagccttggg aacaggaggg ctttgataact cttggcacaaa 60  
atgggaagaa tttttttttt ctcctgcctg aaggacatgt atgatttcgg attcccccaag 120  
gagggtttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atccagcaga ctttcaatct ttccagcaca aaggactcat ctgctacttg ggaacagagc 240  
ctccttagaaa aattttccac tgaactttac cagcaactgta ataacctgga agcatgtgt 300  
15   atacaggagg ttgggggttgg aagagactccc ctgatgaatg aggactccat cttggctgtg 360  
aggaaatact ttccaaagaat cacttttat ctgatggaga agaaatacag cccttgc 420  
tgggagggtt atggcaacca aatcatgaga tttttttttt tttcaacaaa cttgcaaaaa 480  
498  
agattaagga ggaaggaa

20   <210> 25  
<211> 498  
<212> DNA  
<213> Artificial Sequence

25   <220>

<223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID 2C3

30   <400> 25

35   tttttttttt ctcagaccca cagccttggg aacaggaggg ctttgataact cttggcacaaa 60  
atgggaagaa tttttttttt ctcctgcctg aaggacatgt atgactttgg atttcctcag 120  
gagggtttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atccagcaga ctttcaatct ttccagcaca aaggactcat ctgataacttg ggatgcgacc 240  
ctttagaaa aattttccac tgaactttac cagcagctgta atgacctgga agcctgcgtg 300  
40   atacaggagg ttgggggttgg aagagaccccc ctgatgaatg tggactccat cttggctgtg 360  
aggaaatact ttccaaagaat cacttttat ctgacagagaa agaaatacag cccttgc 420  
tgggagggtt atggcaacca aatcatgaga tttttttttt tttcaacaaa cttgcaaaaa 480  
498  
agattaagga ggaaggaa

<210> 26  
<211> 498  
<212> DNA  
45   <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

50   <220>

<223> Clone ID 2D1

<400> 26

55   tttttttttt ctcagaccca cagccttggg aacaggaggg ctttgataact cttggcacaaa 60  
atgggacgaa tttttttttt ctcctgcctg aaggacatgt aagactttgg attcccccaag 120  
gagggtttt atggcaacccg gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atccagcaga ctttcaatct ttccagcaca aagaactcat ctgctgcgtg ggaacagagc 240  
ctccttagaaa aattttccac tgaactttac cagcagctgta atgacctgga agcctgcgtg 300  
atacaggagg ttgggggttgg aagagaccccc ctgatgaatg aggactccat cttggctgtg 360

aagaaaatact tccaaagaat cactcttat ctaatagaga gaaaaatacag cccttgc 420  
tgggaggttg tcagagcaga aatcatgaga tcttcctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

5 <210> 27  
<211> 498  
<212> DNA  
<213> Artificial Sequence

10 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2D10

15 <400> 27  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg cttgtatact cttggcacaa 60  
atggaaagag tcttccttt ctccgtctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
20 atccagcaga cttcaatct ttcaagcaca aaggactcat ctgctacttg ggaacagagc 240  
ctcttagaaa aattttccac tgaactttac cagcaactga ataacctgga agcctgcgtg 300  
atacaggagg ttgggggttga agagactccc ctgatgaatg tggactccat cttggctgtg 360  
aagaaaatact tccgaagaat cactcttat ctgacagaga agaaaatacag cccttgc 420  
tgggaggttg tcagagcaga aatcatgaga tcttcctt tttcaacaaa cttgcaaaaa 480  
25 agattaagga ggaaggaa 498

<210> 28  
<211> 498  
<212> DNA  
30 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

35 <220>  
<223> Clone ID 2D7

<400> 28  
tgtgatctgc ctcagaccca cagccttgg aacaggcg 60  
40 atggaaagaa tcttccttt ctccgtctg aaggacagac atgacttcag atttccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgcctt ccatgagatg 180  
atccagcaga cttcaatct ttcaagcaca aaggactcat ctgctacttg ggaacagagc 240  
ctcttagaaa aattttccac tgaactttac cagcaactga ataacctgga agcctgcgtg 300  
atacaggagg ttgggggttga agagactccc ctgatgaatg tggactctat cttggctgtg 360  
45 aagaaaatact tccaaagaat cactcttat ctgacagaga gaaaatacag cccttgc 420  
tgggaggttg tcagagcaga aatcatgaga tcttcctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 29  
50 <211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
55 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2D9

<400> 29  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ccttgatact cctggcacaa 60  
atgggaagaa tctctccccc ctccgcctg aaggacagac atgactttgg attccccccag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
5 atccagcaga ctttcaatct cttcagcaca aaggactcat ctgctacttg ggaacagagc 240  
ctcctagaaa aattttccac tgaacttaac cagcagctga atgaccttgg agcctgcgtg 300  
atacaggagg ttgggggttgg aagactccc ctggtgaatg tggactccat cctggctgtg 360  
aagaataact tccaaagaat cacttttat ctgacagaga agaaatacag cccttgcgc 420  
tgggaggttgc tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
10 agattaagga ggaaggaa 498

<210> 30  
<211> 498  
<212> DNA  
15 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

20 <220>  
<223> Clone ID 2DA2

<400> 30  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ccttgatact cctggcacaa 60  
25 atgggaagaa tctctccccc ctccgcctg aaggacagac aggacttcgg attccccccag 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atgcagcaga ctttcaatct cttcagcaca aagaactcat ctgctgttgg ggaacagagc 240  
ctcctagaaa aattttccac tgaactccac cagcaactga atgaaacttgg agcatgtgtg 300  
atacaggagg ttgggggttgg aagactccc ctgatgaatg tggactccat cctggctgtg 360  
30 aagaataact tccaaagaat cacttttat ctaatagaga ggaataacag cccttgcgc 420  
tgggaggttgc tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

<210> 31  
35 <211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
40 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2DH9

45 <400> 31  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ccttgatgt cctggcacaa 60  
atgggaccaa tctctccccc ctccgcctg aaggacagat atgatttcgg attccccccag 120  
ggggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atgcagcaga ctttcaatct cttcagcaca aaggattcat ctgctgttgg ggaacagagc 240  
50 ctcctagaaa aattttccac tgaactctac cggcagctga atgaccttgg agcctgtgtg 300  
atacaggagg ttgggggttgg aagacccccc ctgatgaatg tggactccat cctggctgtg 360  
aggaagtact tccaaagaat cacttttat ctgacagaga agaagcatag cccttgcgc 420  
tgggaggttgc tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

55 <210> 32  
<211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

5 <220>  
<223> Clone ID 2G11

<400> 32  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cttggcacaa 60  
10 atggaaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag actcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ctttcaatct cttagcaca aaggactcat ctgataacttgg gaaacagaggc 240  
ctcctagaaa aattctacat tgaactttt cagcagctga atgaccttgg agcctgcgtg 300  
15 atacaggagg ttgggttgg aagagactccc ctgatgaatg tggactccat cttggctgtg 360  
agaaaataact tccaaagaat cactcttat ctgacagagg agaaaatacag cccttgcgc 420  
tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
498  
20 <210> 33  
<211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
25 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID 2G12

30 <400> 33  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ctttgatgct catggcacaa 60  
atgaggagaa tctctccccc ccccccgcctg aaggacagat atgatttcgg attccccca 120  
gagggtttt atggcaacca gttccagaag gctcaagcta tcttcctt ccatgagatg 180  
atccagcaga ctttcaatct cttagcaca aagaactcat ctgctgttgg ggtatggagacc 240  
35 ctcctagaca aattctacac tgaactctac cagcagctga atgacttgg agcctgttg 300  
atcgaggagg ggagggttgg agaaaactccc ctgatgaatg cggactccat cttggctgtg 360  
aaggaaataact tccgaagaat cactcttat ctgacagaga agaaaatacag cccttgcgc 420  
tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
498  
40 <210> 34  
<211> 498  
<212> DNA  
<213> Artificial Sequence

45 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
50 <223> Clone ID 2H9

<400> 34  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cttggcacaa 60  
55 atggaaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccagcaga ctttcaatct cttagcaca aaggactcat ctgctacttgg gaaacagaggc 240  
ctcctagaaa aattttccac tgaacttaac cagcagctga atgaccttgg agcctgttg 300  
acacaggagg ttgggttgg aagagactccc ctgatgaatg aggactctat cttggctgtg 360  
aaggaaataact tccaaagaat cactcttat ctgacagaga agaaaatacag cccttgcgc 420

tgggagggttgcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
 agattaagga ggaaggaa 498

5 <210> 35  
 <211> 498  
 <212> DNA  
 <213> Artificial Sequence

10 <220>  
 <223> Description of Artificial Sequence: Synthetic DNA

<220>  
 <223> Clone ID 6BC11

15 <400> 35  
 tgtgatctgc ctcagaccca cagccttggtaacaggaggg cttgtataact cttggcacaaa 60  
 atgggaagaa tctctccccc ttcctgcctg aaggacat atgatttcgg attccccccag 120  
 gaggagtttgcggcaacca gctccagaag gctcaagcca tctctgtcct ccatgagatg 180  
 atcccgacca cttcaatct cttcagcaca aaggattcat ctgtgtttt ggaacagagc 240  
 20 ctccttagaaa aattttccac tgaactaac cagcagctga atgacatttggaa agcctgcgtg 300  
 atacaggagg ttggagtgaa agagactccc ctgatgtatg tggactccat cttggctgtg 360  
 aagaataact tccaaagaat cacttttat ctgacagaga gaaataacag cccttgtgcc 420  
 tgggagggttgcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
 agattaagga ggaaggaa 498

25 <210> 36  
 <211> 166  
 <212> PRT  
 <213> Artificial Sequence

30 <220>  
 <223> Description of Artificial Sequence: Synthetic amino acid

<220>  
 <223> Clone ID 2DH12

<400> 36  
 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Met  
 1 5 10 .  
 40 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
 20 25 30

Arg Gln Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
 45 35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
 50 55 60

50 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Thr  
 65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
 85 90 95

55 Glu Ala Cys Val Ile Gln Glu Val Gly Val Lys Glu Thr Pro Leu Met  
 100 105 110

Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr

115                    120                    125

Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                    135                    140

5                      Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                    155                    160

10                     Arg Leu Arg Arg Lys Glu  
165

<210> 37  
<211> 166  
15                    <212> PRT  
<213> Artificial Sequence

<220>  
20                    <223> Description of Artificial Sequence: Synthetic amino acid  
<220>  
<223> Clone ID 2CA3

<400> 37  
25                    Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asp Arg Arg Ala Met Ile  
1                    5                        10                    15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20                    25                        30

30                    Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35                    40                        45

35                    Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50                    55                        60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65                    70                        75                    80

40                    Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Glu Leu  
85                    90                        95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Gly Glu Thr Pro Leu Met  
100                    105                        110

45                    Asn Gly Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115                    120                        125

50                    Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                    135                        140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                        155                    160

55                    Arg Leu Arg Arg Lys Glu  
165

&lt;210&gt; 38

<211> 166  
<212> PRT  
<213> Artificial Sequence

5    <220>  
      <223> Description of Artificial Sequence: Synthetic amino acid  
  
      <220>  
      <223> Clone ID 4AB9  
10    <400> 38  
      Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
            1                  5                  10                  15  
  
15    Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
            20                  25                  30  
  
      Arg His Asp Phe Gly Phe Pro Arg Glu Glu Phe Asp Gly Asn Gln Phe  
            35                  40                  45  
20    Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Met Gln Gln Thr  
            50                  55                  60  
  
      Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Asp Glu Thr  
            65                  70                  75                  80  
  
      Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Glu Leu  
            85                  90                  95  
  
30    Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
            100                105                110  
  
      Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
            115                120                125  
35    Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ser Trp Glu Val Val  
            130                135                140  
  
      Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
            145                150                155                160  
  
      Arg Leu Arg Arg Lys Glu  
            165  
  
45    <210> 39  
      <211> 166  
      <212> PRT  
      <213> Artificial Sequence  
50    <220>  
      <223> Description of Artificial Sequence: Synthetic amino acid  
  
      <220>  
      <223> Clone ID 2DA4  
  
      <400> 39  
      Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Met  
            1                  5                  10                  15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

5 Arg Gln Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Ser Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Met Gln Gln Thr  
50 55 60

10 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
15 85 90 95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

20 Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

25 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

30 Arg Leu Arg Arg Lys Glu  
165

<210> 40  
<211> 166  
35 <212> PRT  
<213> Artificial Sequence

<220>  
40 <223> Description of Artificial Sequence: Synthetic amino acid  
<220>  
45 <223> Clone ID 3DA11

<400> 40  
45 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Val  
1 5 10 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
50 20 25 30

50 Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
55 50 55 60

55 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

5 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

10 Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

15 Arg Leu Arg Arg Lys Glu  
165

20 <210> 41  
<211> 166  
<212> PRT  
<213> Artificial Sequence

25 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 2DB11

30 <400> 41  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Met  
1 5 10 15

35 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

40 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

45 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

50 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

55 Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys

145                    150                    155                    160  
Arg Leu Arg Arg Lys Glu  
165

5

<210> 42  
<211> 166  
<212> PRT

10 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

15 <220>  
<223> Clone ID 2CA5

<400> 42  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
20        1              5                  10                 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
25        20              25                  30

25 Arg Gln Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Arg Phe  
35        35              40                  45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50        50              55                  60

30 Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Glu Gln Ser  
65        65              70                  75                 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
35        85              90                  95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100       100              105                  110

40 Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115       115              120                  125

Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
45        130              135                  140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145       145              150                  155                 160

50 Arg Leu Arg Arg Lys Glu  
55        165

<210> 43  
<211> 166  
<212> PRT

<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 2G6

5   <400> 43  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1              5                         10                  15

10   Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
      20                         25                         30

      Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
      35                         40                         45

15   Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
      50                         55                         60

      Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
      65                         70                         75                 80

20   Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
      85                         90                         95

25   Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
      100                         105                         110

      Asn Val Asp Pro Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
      115                         120                         125

30   Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
      130                         135                         140

      Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
      145                         150                         155                 160

35   Arg Leu Arg Arg Lys Glu  
      165

40   <210> 44  
<211> 166  
<212> PRT  
<213> Artificial Sequence

45   <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

      <220>  
<223> Clone ID 3AH7

50   <400> 44  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1              5                         10                  15

55   Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
      20                         25                         30

      Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Ser Asn Gln Phe  
      35                         40                         45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

5 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu His Gln Gln Leu Asn Glu Leu  
85 90 95

10 Glu Ala Cys Val Val Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Leu Gln Arg Ile Thr  
15 115 120 125

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

20 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
25 165

<210> 45  
<211> 166  
<212> PRT  
30 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

35 <220>  
<223> Clone ID 2G5

<400> 45  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Met  
40 1 5 10 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

45 Arg Gln Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

50 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
55 85 90 95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

5 Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

10 Arg Leu Arg Arg Lys Glu  
165

15 <210> 46  
<211> 166  
<212> PRT  
<213> Artificial Sequence

20 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 2BA8

25 <400> 46  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

30 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

35 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65 70 75 80

40 Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
45 100 105 110

Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

50 Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

55 Arg Leu Arg Arg Lys Glu  
165

<210> 47  
<211> 166  
<212> PRT  
<213> Artificial Sequence  
5  
<220>  
<223> Description of Artificial Sequence: Synthetic amino acid  
  
<220>  
10 <223> Clone ID 1F3  
  
<400> 47  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15  
15 Leu Leu Gly Gln Met Gly Arg Ile Ser His Phe Ser Cys Leu Lys Asp  
20 25 30  
20 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45  
Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60  
25 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Val Ala Trp Asp Glu Arg  
65 70 75 80  
Leu Leu Asp Lys Leu Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95  
30 Glu Ala Cys Val Met Gln Glu Val Trp Val Gly Gly Thr Pro Leu Met  
100 105 110  
35 Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125  
Leu Tyr Leu Thr Glu Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140  
40 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160  
Arg Leu Arg Arg Lys Glu  
165  
45  
<210> 48  
<211> 166  
<212> PRT  
50 <213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Synthetic amino acid  
55 <220>  
<223> Clone ID 4BE10  
  
<400> 48  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile

1                   5                   10                   15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20   25                                     30

5                   Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
   35                                     40                                     45

10                  Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Ile Met Gln Gln Thr  
   50                                     55                                     60

Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Asp Glu Thr  
65   70                                     75                                     80

15                  Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Glu Leu  
   85                                     90                                     95

Glu Ala Cys Val Ile Gln Gly Val Gly Val Glu Glu Thr Pro Leu Met  
100   105                                     110

20                  Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
   115                                     120                                     125

25                  Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ser Trp Glu Val Val  
   130                                     135                                     140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145   150                                     155                                     160

30                  Arg Leu Arg Arg Lys Glu  
   165

35                  <210> 49  
                        <211> 166  
                        <212> PRT  
                        <213> Artificial Sequence

40                  <220>  
                        <223> Description of Artificial Sequence: Synthetic amino acid

45                  <220>  
                        <223> Clone ID 2DD9

50                  <400> 49  
                        Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Met  
   1                                     5                                     10                                     15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
55   20                                     25                                     30

Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
65   35                                     40                                     45

55                  Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
   50                                     55                                     60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65   70                                     75                                     80

Leu Leu Glu Lys Phe Ser Thr Gly Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

5 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

10 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ser Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
15 145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

20 <210> 50  
<211> 166  
<212> PRT  
<213> Artificial Sequence

25 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>

30 <223> Clone ID 3CA1

<400> 50  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

35 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Gly Leu Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
40 35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

45 Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu  
85 90 95

50 Glu Ala Cys Val Ile Gln Glu Val Gly Met Glu Glu Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
55 115 120 125

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

5 Arg Leu Arg Arg Lys Glu  
165

10 <210> 51  
<211> 166  
<212> PRT  
<213> Artificial Sequence

15 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

20 <220>  
<223> Clone ID 2F8

25 <400> 51  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

30 <220>  
Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

35 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Met Gln Gln Thr  
50 55 60

40 Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

45 Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Glu Leu  
85 90 95

50 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

55 Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

60 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ser Trp Glu Val Val  
130 135 140

65 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

70 Arg Leu Arg Arg Lys Glu  
165

75 <210> 52  
<211> 166  
<212> PRT  
<213> Artificial Sequence

80 <220>

<223> Description of Artificial Sequence: Synthetic amino acid

<220>

<223> Clone ID 6CG3

5

<400> 52

Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Lys Arg Ala Met Met  
1 5 10 15

10

Leu Leu Ala Gln Met Gly Arg Thr Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

15

Gln Arg Ala Gln Ala Ile Phe Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

20

Phe Asn Phe Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
85 90 95

25

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

30

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

35

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

40

<210> 53

<211> 166

<212> PRT

<213> Artificial Sequence

45

<220>

<223> Description of Artificial Sequence: Synthetic amino acid

<220>

50 <223> Clone ID 3CG7

<400> 53

Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Ser Arg Ala Leu Met  
1 5 10 15

55

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe

35                    40                    45

Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met Ile Gln Gln Thr  
50                    55                    60

5 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Asn  
65                    70                    75                    80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu  
10                    85                    90                    95

Glu Ala Cys Val Ile Gln Glu Val Gly Met Glu Glu Thr Pro Leu Met  
100                  105                  110

15 Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115                  120                  125

Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                  135                  140

20 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                  150                  155                  160

Arg Leu Arg Arg Lys Glu  
25                  165

<210> 54  
30 <211> 166  
<212> PRT  
<213> Artificial Sequence

<220>  
35 <223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 1D3

<400> 54  
40 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1                  5                  10                  15

Leu Leu Ala Gln Met Gly Arg Ile Ser His Phe Ser Cys Leu Lys Asp  
45                  20                  25                  30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly His Gln Phe  
35                  40                  45

Gln Lys Thr Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50                  55                  60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65                  70                  75                  80

55 Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85                  90                  95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100                  105                  110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

5 Leu Tyr Leu Met Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

10 Arg Leu Arg Arg Lys Glu  
165

15 <210> 55  
<211> 166  
<212> PRT  
<213> Artificial Sequence

20 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 2G4

25 <400> 55  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Met Met  
1 5 10 15

30 Leu Leu Ala Gln Met Ser Arg Ile Ser Pro Ser Ser Cys Leu Met Asp  
20 25 30

Arg His Asp Phe Glu Pro Gln Glu Glu Phe Asp Asp Lys Gln Phe  
35 40 45

35 Gln Lys Ala Pro Ala Ile Ser Val Leu His Glu Val Ile Gln Gln Thr  
50 55 60

40 Phe Asn Leu Phe Ser Thr Glu Asp Ser Ser Ala Ala Trp Glu Gln Thr  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

45 Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu Thr Pro Leu Met  
100 105 110

Asn Ala Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

50 Leu Tyr Leu Thr Lys Lys Lys Tyr Ser Pro Cys Ser Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

5       <210> 56  
      <211> 166  
      <212> PRT  
      5    <213> Artificial Sequence

10      <220>  
      <223> Description of Artificial Sequence: Synthetic amino acid

15      10    <220>  
      <223> Clone ID 1A1

20      15    <400> 56  
      Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1               5               10               15

25      20    Leu Leu Ala Gln Met Gly Arg Ile Ser His Phe Ser Cys Leu Lys Asp  
      20               25               30

30      25    Arg Tyr Asp Phe Gly Phe Pro Gln Glu Val Phe Asp Gly Asn Gln Phe  
      35               40               45

35      30    Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met Met Gln Gln Thr  
      50               55               60

40      35    Phe Asn Leu Phe Ser Thr Glu Asp Ser Ser Ala Ala Trp Glu Gln Ser  
      65               70               75               80

45      40    Leu Leu Glu Lys Phe Ser Thr Glu Leu His Gln Gln Leu Asn Asp Leu  
      85               90               95

50      45    Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
      100             105             110

55      50    Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
      115             120             125

60      55    Leu Tyr Leu Met Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
      130             135             140

65      60    Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
      145             150             155             160

70      65    Arg Leu Arg Arg Lys Glu  
      165

75      70    <210> 57  
      <211> 166  
      <212> PRT  
      50    <213> Artificial Sequence

80      75    <220>  
      <223> Description of Artificial Sequence: Synthetic amino acid

85      80    <220>  
      <223> Clone ID 1D10

90      85    <400> 57

Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

5 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Arg Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Leu  
35 40 45

10 Gln Lys Thr Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
65 70 75 80

15 Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
85 90 95

Glu Ala Cys Val Ile Gln Gly Val Gly Val Glu Glu Thr Pro Pro Met  
20 100 105 110

Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

25 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

30 Arg Leu Arg Arg Lys Glu  
165

35 <210> 58  
<211> 166  
<212> PRT  
<213> Artificial Sequence

40 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 1F6

45 <400> 58  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Thr Leu Met  
1 5 10 15

50 Ile Met Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

55 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser

	65	70	75	80
	Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu			
	85	90		95
5	Glu Ala Cys Val Ile Gln Glu Ala Gly Val Glu Glu Thr Pro Leu Met			
	100	105		110
10	Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr			
	115	120	125	
	Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val			
	130	135	140	
15	Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys			
	145	150	155	160
	Arg Leu Arg Arg Lys Glu			
	165			
20	<210> 59			
	<211> 166			
	<212> PRT			
25	<213> Artificial Sequence			
	<220>			
	<223> Description of Artificial Sequence: Synthetic amino acid			
30	<220>			
	<223> Clone ID 2A10			
	<400> 59			
35	Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile			
	1	5	10	15
	Leu Leu Ala Gln Met Gly Arg Ile Ser His Phe Ser Cys Leu Lys Asp			
	20	25	30	
40	Arg Tyr Asp Phe Gly Phe Pro Gln Glu Val Phe Asp Gly Asn Gln Phe			
	35	40	45	
	Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met Ile Gln Gln Thr			
	50	55	60	
45	Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser			
	65	70	75	80
50	Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu			
	85	90	95	
	Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met			
	100	105	110	
55	Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr			
	115	120	125	
	Leu Tyr Leu Met Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val			
	130	135	140	

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

5 Arg Leu Arg Arg Lys Glu  
165

10 <210> 60  
<211> 166  
<212> PRT  
<213> Artificial Sequence

15 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

20 <220>  
<223> Clone ID 2C3

25 <400> 60  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
25 20 25 30

30 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Ser  
35 35 40 45

35 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

40 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Asp Thr Trp Asp Ala Thr  
65 70 75 80

45 Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
85 90 95

50 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

55 Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

60 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

65 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

70 Arg Leu Arg Arg Lys Glu  
165

75 <210> 61  
<211> 166  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

5    <220>  
    <223> Clone ID 2D1

<400> 61  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1              5                          10                        15

10    Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
      20    25                                30

15    Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
      35    40                                45

Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met Ile Gln Gln Thr  
      50    55                                60

20    Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
      65    70                                75                                80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu  
      85    90                                        95

25    Glu Ala Cys Val Ile Gln Glu Val Gly Met Glu Glu Thr Pro Leu Met  
      100    105                                110

30    Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
      115    120                                125

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
      130    135                                        140

35    Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
      145    150                                155                                160

Arg Leu Arg Arg Lys Glu  
      165

40

<210> 62  
<211> 166  
<212> PRT  
45    <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

50    <220>  
    <223> Clone ID 2D10

<400> 62  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1              5                          10                        15

55    Leu Leu Ala Gln Met Gly Arg Val Ser Pro Phe Ser Cys Leu Lys Asp  
      20    25                                30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
 35 40 45  
 Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met Ile Gln Gln Thr  
 5 50 55 60  
 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
 65 70 75 80  
 10 Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu  
 85 90 95  
 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
 15 100 105 110  
 Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
 115 120 125  
 20 Leu Tyr Leu Thr Glu Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
 130 135 140  
 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
 25 145 150 155 160  
 Arg Leu Arg Arg Lys Glu  
 165  
 30 <210> 63  
 <211> 166  
 <212> PRT  
 <213> Artificial Sequence  
 35 <220>  
 <223> Description of Artificial Sequence: Synthetic amino acid  
 <220>  
 <223> Clone ID 2D7  
 40 <400> 63  
 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
 1 5 10 15  
 45 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
 20 25 30  
 Arg His Asp Phe Arg Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
 35 40 45  
 50 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
 50 55 60  
 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
 55 65 70 75 80  
 Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu Asn Asn Leu  
 85 90 95  
 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met

	100	105	110
	Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr		
	115	120	125
5	Leu Tyr Leu Thr Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val		
	130	135	140
10	Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys		
	145	150	155
	Arg Leu Arg Arg Lys Glu		
	165		
15	<210> 64		
	<211> 166		
	<212> PRT		
	<213> Artificial Sequence		
20	<220>		
	<223> Description of Artificial Sequence: Synthetic amino acid		
	<220>		
25	<223> Clone ID 2D9		
	<400> 64		
	Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile		
	1	5	10
30			15
	Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp		
	20	25	30
	Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe		
35	35	40	45
	Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr		
	50	55	60
40	Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser		
	65	70	75
			80
	Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu		
	85	90	95
45	Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Val		
	100	105	110
	Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr		
50	115	120	125
	Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val		
	130	135	140
55	Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys		
	145	150	155
	Arg Leu Arg Arg Lys Glu		
	165		

5           <210> 65  
          <211> 166  
          <212> PRT  
          <213> Artificial Sequence

10          <220>  
          <223> Description of Artificial Sequence: Synthetic amino acid

10          <220>  
          <223> Clone ID 2DA2

15          <400> 65  
          Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Pro Leu Ile  
              1                   5                   10                   15

20          Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
              20                   25                   30

20          Arg Gln Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
              35                   40                   45

25          Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Met Gln Gln Thr  
              50                   55                   60

30          Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Glu Gln Ser  
              65                   70                   75                   80

35          Leu Leu Glu Lys Phe Ser Thr Glu Leu His Gln Gln Leu Asn Glu Leu  
              85                   90                   95

35          Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
              100                   105                   110

40          Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
              115                   120                   125

40          Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
              130                   135                   140

45          Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
              145                   150                   155                   160

45          Arg Leu Arg Arg Lys Glu  
              165

50          <210> 66  
          <211> 166  
          <212> PRT  
          <213> Artificial Sequence

55          <220>  
          <223> Description of Artificial Sequence: Synthetic amino acid

55          <220>  
          <223> Clone ID 2DH9

<400> 66  
Cys Asp Leu Pro Gln Thr His Ser Pro Gly Asn Arg Arg Ala Leu Met  
1 5 10 15

5 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg Tyr Asp Phe Gly Phe Pro Gln Gly Glu Phe Asp Gly Asn Gln Phe  
35 40 45

10 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Met Gln Gln Thr  
50 55 60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
15 65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Arg Gln Leu Asn Asp Leu  
85 90 95

20 Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

25 Leu Tyr Leu Thr Glu Lys Lys His Ser Pro Cys Ser Trp Glu Val Val  
130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
30 145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

35 <210> 67  
<211> 166  
<212> PRT  
<213> Artificial Sequence

40 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

45 <220>  
<223> Clone ID 2G11

<400> 67  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

50 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

Arg His Asp Phe Gly Leu Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
55 35 40 45

Gln Lys Thr Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Asp Thr Trp Glu Gln Ser  
65 70 75 80

5 Leu Leu Glu Lys Phe Tyr Ile Glu Leu Phe Gln Gln Leu Asn Asp Leu  
85 90 95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

10 Asn Val Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

Leu Tyr Leu Thr Glu Glu Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

15 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
20 165

<210> 68  
<211> 166  
25 <212> PRT  
<213> Artificial Sequence

<220>  
30 <223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID 2G12

<400> 68  
35 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Thr Leu Met  
1 5 10 15

Leu Met Ala Gln Met Arg Arg Ile Ser Pro Phe Pro Arg Leu Lys Asp  
20 25 30

40 Arg Tyr Asp Phe Gly Phe Pro Gln Glu Val Phe Asp Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Phe Leu Phe His Glu Met Met Gln Gln Thr  
45 50 55 60

Phe Asn Leu Phe Ser Thr Lys Asn Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

50 Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

Glu Ala Cys Val Met Gln Glu Gly Arg Val Gly Glu Thr Pro Leu Met  
100 105 110

55 Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
115 120 125

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Ala Val

130                    135                    140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                    155                    160

5                    Arg Leu Arg Arg Lys Glu  
                          165

10 <210> 69  
<211> 166  
<212> PRT  
<213> Artificial Sequence

15 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid  
  
<220>  
<223> Clone ID 2H9

20 <400> 69  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1                    5                    10                    15

25 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20                    25                    30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35                    40                    45

30 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50                    55                    60

35 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
65                    70                    75                    80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
85                    90                    95

40 Glu Ala Cys Val Thr Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100                    105                    110

Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115                    120                    125

45 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                    135                    140

50 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                    155                    160

Arg Leu Arg Arg Lys Glu  
165

55 <210> 70  
<211> 166  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

5 <220>  
<223> Clone ID 6BC11

<400> 70  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
10 1 5 10 15

Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 20 25 30

15 Arg Tyr Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Leu  
35 35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

20 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Glu Gln Ser  
65 70 75 80

Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
25 85 90 95

Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

30 Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr  
115 120 125

Leu Tyr Leu Thr Glu Arg Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

35 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

40 Arg Leu Arg Arg Lys Glu  
165

<210> 71  
<211> 166  
45 <212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic amino acid

50 <220>  
<223> Clone ID t19bb

<220>  
55 <221> MOD\_RES  
<222> (11)  
<223> N or D

<220>

<221> MOD\_RES  
<222> (12)  
<223> R, S, or K

5 <220>  
<221> MOD\_RES  
<222> (15)  
<223> L or M

10 <220>  
<221> MOD\_RES  
<222> (16)  
<223> I, M or V

15 <220>  
<221> MOD\_RES  
<222> (19)  
<223> A or G

20 <220>  
<221> MOD\_RES  
<222> (22)  
<223> G or R

25 <220>  
<221> MOD\_RES  
<222> (24)  
<223> I or T

30 <220>  
<221> MOD\_RES  
<222> (26)  
<223> P or H

35 <220>  
<221> MOD\_RES  
<222> (34)  
<223> H, Y or Q

40 <220>  
<221> MOD\_RES  
<222> (38)  
<223> F or L

45 <220>  
<221> MOD\_RES  
<222> (40)  
<223> Q or R

50 <220>  
<221> MOD\_RES  
<222> (45)  
<223> G or S

55 <220>  
<221> MOD\_RES  
<222> (46)  
<223> N or H

5           <220>  
          <221> MOD\_RES  
          <222> (47)  
          <223> Q or R  
  
10          <220>  
          <221> MOD\_RES  
          <222> (50)  
          <223> K or R  
  
15          <220>  
          <221> MOD\_RES  
          <222> (51)  
          <223> A or T  
  
20          <220>  
          <221> MOD\_RES  
          <222> (55)  
          <223> S or F  
  
25          <220>  
          <221> MOD\_RES  
          <222> (56)  
          <223> V or A  
  
30          <220>  
          <221> MOD\_RES  
          <222> (57)  
          <223> L or F  
  
35          <220>  
          <221> MOD\_RES  
          <222> (60)  
          <223> M or I  
  
40          <220>  
          <221> MOD\_RES  
          <222> (61)  
          <223> I or M  
  
45          <220>  
          <221> MOD\_RES  
          <222> (67)  
          <223> L or F  
  
50          <220>  
          <221> MOD\_RES  
          <222> (72)  
          <223> D or N  
  
55          <220>  
          <221> MOD\_RES  
          <222> (75)  
          <223> A or V  
  
          <220>  
          <221> MOD\_RES  
          <222> (76)  
          <223> A or T

5        <220>  
          <221> MOD\_RES  
          <222> (78)  
5        <223> E or D  
  
10      <220>  
          <221> MOD\_RES  
          <222> (79)  
10      <223> Q or E  
  
15      <220>  
          <221> MOD\_RES  
          <222> (80)  
15      <223> S, R, T, or N  
  
20      <220>  
          <221> MOD\_RES  
          <222> (83)  
20      <223> E or D  
  
25      <220>  
          <221> MOD\_RES  
          <222> (85)  
25      <223> F or L  
  
30      <220>  
          <221> MOD\_RES  
          <222> (86)  
30      <223> S or Y  
  
35      <220>  
          <221> MOD\_RES  
          <222> (88)  
35      <223> E or G  
  
40      <220>  
          <221> MOD\_RES  
          <222> (90)  
40      <223> Y, H, N  
  
45      <220>  
          <221> MOD\_RES  
          <222> (95)  
45      <223> D, E, or N  
  
50      <220>  
          <221> MOD\_RES  
          <222> (101)  
50      <223> I, M, or V  
  
55      <220>  
          <221> MOD\_RES  
          <222> (103)  
55      <223> E or G  
  
55      <220>  
          <221> MOD\_RES  
          <222> (105)

<223> G or W  
5      <220>  
      <221> MOD\_RES  
      <222> (106)  
      <223> V or M  
  
10     <220>  
      <221> MOD\_RES  
      <222> (107)  
      <223> E, G, or K  
  
15     <220>  
      <221> MOD\_RES  
      <222> (108)  
      <223> E or G  
  
20     <220>  
      <221> MOD\_RES  
      <222> (114)  
      <223> V, E, or G  
  
25     <220>  
      <221> MOD\_RES  
      <222> (116)  
      <223> S or P  
  
30     <220>  
      <221> MOD\_RES  
      <222> (121)  
      <223> K or R  
  
35     <220>  
      <221> MOD\_RES  
      <222> (124)  
      <223> F or L  
  
40     <220>  
      <221> MOD\_RES  
      <222> (132)  
      <223> T, I, or M  
  
45     <220>  
      <221> MOD\_RES  
      <222> (134)  
      <223> K or R  
  
50     <220>  
      <221> MOD\_RES  
      <222> (140)  
      <223> A or S  
  
55     <400> 71  
      Cys Asp Leu Pro Gln Thr His Ser Leu Gly Xaa Xaa Arg Ala Xaa Xaa  
                1                  5                  10                  15  
      Leu Leu Xaa Gln Met Xaa Arg Xaa Ser Xaa Phe Ser Cys Leu Lys Asp  
                20                  25                  30

Arg Xaa Asp Phe Gly Xaa Pro Xaa Glu Glu Phe Asp Xaa Xaa Xaa Phe  
35 40 45

5 Gln Xaa Xaa Gln Ala Ile Xaa Xaa Xaa His Glu Xaa Xaa Gln Gln Thr  
50 55 60

Phe Asn Xaa Phe Ser Thr Lys Xaa Ser Ser Xaa Xaa Trp Xaa Xaa Xaa  
65 70 75 80

10 Leu Leu Xaa Lys Xaa Xaa Thr Xaa Leu Xaa Gln Gln Leu Asn Xaa Leu  
85 90 95

Glu Ala Cys Val Xaa Gln Xaa Val Xaa Xaa Xaa Xaa Thr Pro Leu Met  
100 105 110

15 Asn Xaa Asp Xaa Ile Leu Ala Val Xaa Lys Tyr Xaa Gln Arg Ile Thr  
115 120 125

Leu Tyr Leu Xaa Glu Xaa Lys Tyr Ser Pro Cys Xaa Trp Glu Val Val  
20 130 135 140

Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

25 Arg Leu Arg Arg Lys Glu  
165

30 <210> 72  
<211> 498  
<212> DNA  
<213> Artificial Sequence

35 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID CH1.1

40 <400> 72  
tgcgtatctgc ctcagaccca cagccttggt aacaggagggg cttgtatact cttggcacaa 60  
atggaaagaa tctctccctt ctccgtctg atggacagac atgactttgg atttccccag 120  
gaggagttt gatacaacca gttccagaag gctcaagcca tctctgtcct ccatgagatg 180  
atccaacaga cttcaatct cttagcaca aaggactcat ctgtacttg ggatgagaca 240  
45 cttctagaca aattctacac tgaaccttac cagcagctga atgaccctgga agcctgcgtg 300  
atacaggagg ttgggggtgga agagactccc ctgatgaatg aggactccat ctggctgtg 360  
aagaaatact tccgaagaat cactcttat ctgacagaga agaaatacag cccttgtgcc 420  
tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

50 <210> 73  
<211> 498  
<212> DNA  
<213> Artificial Sequence

55 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>

<223> Clone ID CH1.2

5       <400> 73  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ctttgataact cctggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt atggcaaccca gttccagaag gctcaaggca tctctgtccct ccatgagatg 180  
atccagcaga cttccatct cttcagcaca aaggactcat ctgctactt ggaacagagc 240  
ctccttagaaa aattttccac tgaacttaac cagcagctga atgaccttgg agcctgcgtg 300  
atacaggagg ttggggttgg aagagactccc ctgatgaatg tggactccat cctggctgtg 360  
10      aagaaataact tccgaagaat cactcttat ctgacagaga agaaatacag cccttgcgtc 420  
tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
  498

15      <210> 74  
15      <211> 498  
15      <212> DNA  
15      <213> Artificial Sequence

20      <220>  
20      <223> Description of Artificial Sequence: Synthetic DNA

25      <220>  
25      <223> Clone ID CH1.3

30      <400> 74  
tgtgatctgc ctcagaccca cagccttgg aacaggagga ctttgatgat aatggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccctcag 120  
gaggagttt atggcaaccca gttccagaag gctcaaggca tctctgtccct ccatgagatg 180  
atccagcaga cttccatct cttcagcaca aaggactcat ctgctactt ggatgagaca 240  
30      cttctagaca aattctacac tgaactttac cagcagctga atgaccttgg agcctgtatg 300  
atgcaggagg ttggggttgg agacactccct ctgatgaatg tggactctat cctgactgtg 360  
agaaataact tccgaagaat cactcttat ctgacagaga agaaatacag cccttgcgtc 420  
tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
  498

35      <210> 75  
35      <211> 498  
35      <212> DNA  
35      <213> Artificial Sequence

40      <220>  
40      <223> Description of Artificial Sequence: Synthetic DNA

45      <220>  
45      <223> Clone ID CH1.4

50      <400> 75  
tgtgatctgc ctcagaccca cagcctgggt aataggaggg ctttgataact cctggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg attccccca 120  
gaggagttt gtggcaaccca gttccagaag gctcaaggca tctctgtccct ccatgagatg 180  
atccagcaga cttccatct cttcagcaca gaggactcat ctgctgttgg ggatgagacc 240  
50      ctccttagaca aattctacat tgaactttt cagcaacttga atgaccttgg agcctgtgtg 300  
atgcaggagg agagggtggg agaaactccc ctgatgaatg cggactccat ctggctgtg 360  
aagaaataact tccaaagaat cactcttat ctgacagaga agaaatacag cccttgcgtc 420  
55      tggaggttg tcagagcaga aatcatgaga tctttctt tttcaacaaa cttgcaaaaa 480  
  498

  498

<212> DNA  
<213> Artificial Sequence

5 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID CH2.1

10 <400> 76  
tgtgatctgc ctcagaccca cagccttgg aacaggagga ctgtatgtat aatggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg atttcctcg 120  
gaggagttt atggcaacca gttccagaag gctcaagccca tctctgtccct ccatgagatg 180  
atccagcaga cttcaatct cttcagcaca aaggactcat ctgtacttg ggatgagaca 240  
15 cttcttagaca aattctacac tgaactttac cagcagctga atgaccttgg agcctgtatg 300  
atacaggagg ttgggggttgg aagactccc ctgatgaatg aggactccat cttggctgtg 360  
aagaataact tccgaagaat cactcttat ctgacagaga agaaatacag cccttgc 420  
tgggagttt tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcacaaa 480  
498  
20 agattaagga ggaaggaa

<210> 77  
<211> 498  
<212> DNA  
<213> Artificial Sequence

25 <220>  
<223> Description of Artificial Sequence: Synthetic DNA

<220>  
<223> Clone ID CH2.2

30 <400> 77  
tgtgatctgc ctcagaccca cagccttgg aacaggaggg ctgtataact cttggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg atggacagac atgactttgg atttcctcg 120  
35 gaggagttt atgacaacca gttccagaag gctcaagccca tctctgtccct ccatgagatg 180  
atccaaacaga cttcaatct cttcagcaca aaggactcat ctgtacttg ggatgagaca 240  
cttcttagaca aattctacac tgaactttac cagcagctga atgaccttgg agcctgtatg 300  
atgcaggagg ttggagtttgg aagactccat ctgatgaatg tggactctat cttgactgtg 360  
40 aagaataact tccgaagaat cactcttat ctgacagaga agaaatacag cccttgc 420  
tgggagttt tcagagcaga aatcatgaga tctttcttt tttcaacaaa cttgcacaaa 480  
498  
45 <210> 78  
<211> 498  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Synthetic DNA

50 <220>  
<223> Clone ID CH2.3

<400> 78  
tgtgatctgc ctcagaccca cagccttgg aacaggagga ctgtatgtat aatggcacaa 60  
atgggaagaa tctctccccc ctcctgcctg aaggacagac atgactttgg atttcctcg 120  
gaggagttt atggcaacca gttccagaag gctcaagccca tctctgtccct ccatgagatg 180  
atccagcaga cttcaatct cttcagcaca aaggactcat ctgtacttg ggatgagaca 240  
55 cttcttagaca aattctacac tgaactttac cagcagctga atgaccttgg agcctgtatg 300

atgcaggagg ttggagtgga agacactcct ctgatgaatg aggactccat cttggctgtg 360  
aagaataact tccgaagaat cactcttat ctgacagaga agaaaatacag cccttgc 420  
tggaggttg tcagagcaga aatcatgaga tctttcttt tctcaacaaa cttgcaaaaa 480  
agattaagga ggaaggaa 498

5           <210> 79  
          <211> 166  
          <212> PRT  
          <213> Artificial Sequence

10          <220>  
          <223> Description of Artificial Sequence: Synthetic amino acid

15          <220>  
          <223> Clone ID CH1.1

          <400> 79  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
      1                  5                                 10                  15

20          Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Met Asp  
      20                  25                                 30

          Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Asp Asn Gln Phe  
      35                  40                                 45

          Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
      50                  55                                 60

30          Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Asp Glu Thr  
      65                  70                                 75                  80

          Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
      85                  90                                 95

35          Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
      100                 105                                 110

40          Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
      115                 120                                 125

          Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
      130                 135                                 140

45          Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
      145                 150                                 155                  160

          Arg Leu Arg Arg Lys Glu  
                         165

50

          <210> 80  
          <211> 166  
          <212> PRT  
          <213> Artificial Sequence

          <220>  
          <223> Description of Artificial Sequence: Synthetic amino acid

&lt;220&gt;

&lt;223&gt; Clone ID CH1.2

&lt;400&gt; 80

5 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 3010 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 4515 Gln Lys Ala Gln Gly Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60Phe His Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Glu Gln Ser  
65 70 75 8020 Leu Leu Glu Lys Phe Ser Thr Glu Leu Asn Gln Gln Leu Asn Asp Leu  
85 90 95Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 11025 Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
115 120 12530 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 16035 Arg Leu Arg Arg Lys Glu  
165

40 &lt;210&gt; 81

&lt;211&gt; 166

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

45 &lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Synthetic amino acid

&lt;220&gt;

&lt;223&gt; Clone ID CH1.3

50 &lt;400&gt; 81

Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Thr Leu Met  
1 5 10 15Ile Met Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 3055 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

5 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Asp Glu Thr  
65 70 75 80

Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

10 Glu Ala Cys Met Met Gln Glu Val Gly Val Glu Asp Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Thr Val Arg Lys Tyr Phe Arg Arg Ile Thr  
115 120 125

15 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

20 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

25 <210> 82  
<211> 166  
<212> PRT  
<213> Artificial Sequence

30 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

35 <220>  
<223> Clone ID CH1.4

<400> 82  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

40 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

45 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

50 Phe Asn Leu Phe Ser Thr Glu Asp Ser Ser Ala Ala Trp Asp Glu Thr  
65 70 75 80

Leu Leu Asp Lys Phe Tyr Ile Glu Leu Phe Gln Gln Leu Asn Asp Leu  
85 90 95

55 Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu Thr Pro Leu Met  
100 105 110

Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile Thr

115                    120                    125  
Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                    135                    140  
5 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                    155                    160  
10 Arg Leu Arg Arg Lys Glu  
165  
  
<210> 83  
<211> 166  
15 <212> PRT  
<213> Artificial Sequence  
  
<220>  
20 <223> Description of Artificial Sequence: Synthetic amino acid  
<220>  
<223> Clone ID CH2.1  
  
<400> 83  
25 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Thr Leu Met  
1                    5                    10                    15  
  
Ile Met Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20                    25                    30  
30 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35                    40                    45  
  
Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
35                    50                    55                    60  
  
Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Asp Glu Thr  
65                    70                    75                    80  
  
40 Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85                    90                    95  
  
Glu Ala Cys Met Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100                    105                    110  
45 Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
115                    120                    125  
  
50 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130                    135                    140  
Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145                    150                    155                    160  
55 Arg Leu Arg Arg Lys Glu  
165

&lt;210&gt; 84

<211> 166  
<212> PRT  
<213> Artificial Sequence

5 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

<220>  
<223> Clone ID CH2.2

10 <400> 84  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu Ile  
1 5 10 15

15 Leu Leu Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Met Asp  
20 25 30

Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Asp Asn Gln Phe  
35 40 45

20 Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

25 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Asp Glu Thr  
65 70 75 80

Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
85 90 95

30 Glu Ala Cys Met Met Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

Asn Val Asp Ser Ile Leu Thr Val Lys Lys Tyr Phe Arg Arg Ile Thr  
115 120 125

35 Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

40 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
165

45 <210> 85  
<211> 166  
<212> PRT  
<213> Artificial Sequence

50 <220>  
<223> Description of Artificial Sequence: Synthetic amino acid

55 <220>  
<223> Clone ID CH2.3

<400> 85  
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Thr Leu Met  
1 5 10 15

Ile Met Ala Gln Met Gly Arg Ile Ser Pro Phe Ser Cys Leu Lys Asp  
20 25 30

5 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe  
35 40 45

Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln Thr  
50 55 60

10 Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Thr Trp Asp Glu Thr  
65 70 75 80

Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu  
15 85 90 95

Glu Ala Cys Met Met Gln Glu Val Gly Val Glu Glu Thr Pro Leu Met  
100 105 110

20 Asn Glu Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr  
115 120 125

Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val  
130 135 140

25 Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr Asn Leu Gln Lys  
145 150 155 160

Arg Leu Arg Arg Lys Glu  
30 165

<210> 86  
<211> 15  
35 <212> DNA  
<213> Artificial Sequence

<220>  
40 <223> Description of Artificial Sequence: Synthetic DNA

<400> 86  
tgcgacttac caca 15

45 <210> 87  
<211> 26  
<212> PRT  
<213> Artificial Sequence

<220>  
50 <223> Description of Artificial Sequence: Synthetic amino acid

<400> 87  
Trp Glu Val Val Arg Ser Glu Ile Met Arg Ser Phe Ser Tyr Ser Thr  
1 5 10 15

55 Asn Leu Gln Arg Arg Leu Arg Arg Lys Asp  
20 25

<210> 88  
<211> 26  
<212> PRT  
<213> Artificial Sequence  
5  
<220>  
<223> Description of Artificial Sequence: Synthetic amino acid  
  
<400> 88  
10 Trp Glu Leu Val Arg Ala Glu Ile Val Arg Ser Phe Ser Phe Ser Thr  
1 5 10 15  
Asn Leu Asn Lys Arg Leu Arg Lys Lys Glu  
20 25  
15

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number  
**WO 01/025438 A3**

- (51) International Patent Classification<sup>7</sup>: C12N 15/21, C07K 14/56, C12N 1/21, A61K 31/70, C12N 15/62, C07K 16/24, A61K 38/21, C12N 5/10, G06F 19/00
- (74) Agents: QUINE, Jonathan, Alan; The Law Offices of Jonathan Alan Quine, P.O. Box 458, Alameda, CA 94501 et al. (US).
- (21) International Application Number: PCT/US00/27781
- (22) International Filing Date: 6 October 2000 (06.10.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/415,183 7 October 1999 (07.10.1999) US
- (71) Applicant (for all designated States except US): MAXY-GEN, INC. [US/US]; 515 Galveston Drive, Redwood City, CA 94063 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): HEINRICH, Volker [DE/US]; Apartment 9, 1915 Mount Vernon Court, Mountain View, CA 94040 (US). CHEN, Teddy [US/US]; 3652 McNulty Way, Redwood City, CA (US). PATTEN, Phillip, A. [US/US]; 261 La Cuesta Drive, Menlo Park, CA 94025 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
11 July 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/025438 A3

(54) Title: IFN-ALPHA HOMOLOGUES

(57) Abstract: Alpha interferon homologues (both nucleic acids and polypeptides) are provided. Compositions including these interferon homologue polypeptides and nucleic acids, recombinant cells comprising said homologue polypeptides and nucleic acids, methods of making the new homologues, antibodies to the new homologues, and methods of using the homologues are provided. Integrated systems comprising the sequences of the nucleic acids or polypeptides are also provided.

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/27781

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>					
IPC 7 C12N15/21 C07K14/56 C12N1/21 A61K31/70 C12N15/62 C07K16/24 A61K38/21 C12N5/10 G06F19/00					
According to International Patent Classification (IPC) or to both national classification and IPC					
<b>B. FIELDS SEARCHED</b>					
Minimum documentation searched (classification system followed by classification symbols)					
IPC 7 C12N C07K A61K G06F					
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched					
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)					
EPO-Internal, SEQUENCE SEARCH, BIOSIS, WPI Data, EMBL					
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>					
Category *	Citation of document, with indication, where appropriate, of the relevant passages				Relevant to claim No.
X	CHANG C -C J ET AL: "EVOLUTION OF A CYTOKINE USING DNA FAMILY SHUFFLING" NATURE BIOTECHNOLOGY, NATURE PUBLISHING, US, vol. 17, no. 8, August 1999 (1999-08), pages 793-797, XP000946490 ISSN: 1087-0156 cited in the application				1,4, 8-11, 20-27, 31,60, 61, 66-71, 76, 81-87, 89,90, 104-123
A	the whole document especially page 794; figure 1B  page 794; table 1 page 796, left-hand column, paragraph 2 -/-/				33-35, 37, 41-45, 51,63, 65,93, 94,101
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.			<input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed					
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family					
Date of the actual completion of the international search			Date of mailing of the international search report		
29 August 2001			15.01.02		
Name and mailing address of the ISA			Authorized officer		
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016			LE CORNEC N.D.R.		

## INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/27781

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>---</p> <p>SCAROZZA A M ET AL: "DNA SYNTHESIS IN NUCLEI ISOLATED FROM DAUDI B CELLS: A MODEL TO STUDY THE ANTIPROLIFERATIVE MECHANISMS OF INTERFERON-ALPHA" JOURNAL OF INTERFERON RESEARCH, MARY ANN LIEBERT, INC., NEW YORK, NY, US, vol. 12, 1992, pages 35-42, XP001002244 ISSN: 0197-8357 cited in the application the whole document</p> <p>---</p>	33,35,53
X	<p>WO 83 02457 A (CETUS CORP) 21 July 1983 (1983-07-21) the whole document</p> <p>---</p>	1,20-25
X	<p>DATABASE EMBL [Online] HSIFNF, accession number M28586, 26 November 1990 (1990-11-26)</p> <p>E. GREN ET AL: "human leukocyte interferon-alpha mRNA, complete cds, clone pIFN9" XP002176121 abstract</p> <p>-&amp; E. GREN ET AL: "Novel leukocyte interferon subtype and structural comparison of alpha interferon genes" JOURNAL OF INTERFERON RESEARCH, vol. 4, 1984, pages 609-617, XP001016190</p> <p>---</p>	1
A	<p>CRAMERI A ET AL: "DNA SHUFFLING OF A FAMILY OF GENES FROM DIVERSE SPECIES ACCELERATES DIRECTED EVOLUTION" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 391, 15 January 1998 (1998-01-15), pages 288-291, XP000775869 ISSN: 0028-0836</p> <p>---</p>	
A	<p>HENCO K ET AL: "STRUCTURAL RELATIONSHIP OF HUMAN INTERFERON ALPHA GENES AND PSEUDOGENES" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 185, no. 2, 20 September 1985 (1985-09-20), pages 227-260, XP000605295 ISSN: 0022-2836 cited in the application</p> <p>---</p> <p>-/-</p>	

## INTERNATIONAL SEARCH REPORT

International Application No PCT/US 00/27781
---

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	D.V. GOEDDEL ET AL: "The structure of eight distinct cloned human leukocyte interferon cDNAs" NATURE, vol. 290, no. 5, 5 March 1981 (1981-03-05), pages 20-26, XP002055149 the whole document ---	
P,X	WO 00 52153 A (MAXYGEN INC ;HOWARD RUSSELL J (US); PATTEN PHILLIP A (US)) 8 September 2000 (2000-09-08)  page 40 -page 52; example 1 -----	1,4, 8-11, 20-27, 31-61, 66-71, 76, 81-87, 89,90, 104-123

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/27781

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  

Although claims 77-80 and claims 81-82, 88 (as far as they concern an *in vivo* method) are directed to a method of treatment of the human/animal body (rule 39.1 (IV) PCT), the search has been carried out and based on the alleged effects of the compound/composition.
2.  Claims Nos.: 68-69, 71-75, 93-103 all partially  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  

see FURTHER INFORMATION sheet PCT/ISA/210
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims (1, 3, 4-123) all partially

#### Remark on Protest

- The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 68-69, 71-75, 93-103 all partially

Claims 68-69 have been searched in view of claim 70.

Claims 71-75 have been searched in view of claim 76.

Although claims 93-103, could be at least partially be considered as a mere presentation of information, rule 39.1 (V) PCT, the search has been carried out as far as possible in our systematic documentation.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1. : Claims (1, 3, 4-123) all partially

Nucleic acid represented by sequence ID no.1 encoding a polypeptide represented by sequence ID no.36. Antibody. Method of production. Therapeutic uses thereof.

Inventions 2. to 35. : Claims (1, 3, 4-123) all partially as far as applicable.

Nucleic acid represented by sequences ID no.2 to no.35 encoding a polypeptide represented respectively by sequences ID no.37 to no.70. Antibody. Method of production. Therapeutic uses thereof.

Invention 36. : claims (2, 20-33, 37-40, 54-123) all partially

Nucleic acid represented by sequence ID no.72 encoding a polypeptide represented by sequence ID no.79. Antibody. Method of production. Therapeutic uses thereof.

Invention 37. to 42. : claims (2, 20-33, 37-40, 54-123) all partially as far as applicable.

Nucleic acid represented by sequences ID no.73 to no.78 encoding a polypeptide represented respectively by sequences ID no.80 to no.85. Antibody. Method of production. Therapeutic uses thereof.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No	
PCT/US 00/27781	

Patent document cited in search report	Publication date		Patent family member(s)	Publication date
WO 8302457	A 21-07-1983	US AU CA EP IT WO	5098703 A 1152783 A 1282356 A1 0098862 A1 1167607 B 8302457 A1	24-03-1992 28-07-1983 02-04-1991 25-01-1984 13-05-1987 21-07-1983
WO 0052153	A 08-09-2000	AU WO	3725600 A 0052153 A2	21-09-2000 08-09-2000

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
6 March 2003 (06.03.2003)

PCT

(10) International Publication Number  
**WO 03/018820 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/864,  
15/62, 15/35, 15/10, 7/01, 5/10, C07K 14/015, C12Q 1/70,  
A61K 39/23, A61P 15/00, 31/18

(21) International Application Number: PCT/IB02/04087

(22) International Filing Date: 16 August 2002 (16.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/315,382 27 August 2001 (27.08.2001) US  
10/022,390 17 December 2001 (17.12.2001) US

(71) Applicant (for all designated States except US): NAUTILUS BIOTECH [FR/FR]; 1, rue Pierre Fontaine,  
F-91000 EVRY (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): VEGA,  
Manuel [IT/FR]; 49, rue Félix Faure, F-91270 VIGNEX-SUR-SEINE (FR). DRITTANTI, Lila [IT/FR];  
49 rue Félix Faure, F-91270 VIGNEX-SUR-SEINE (FR). FLAUX, Marjorie [FR/FR]; 38 rue Jules Vallès,  
F-91000 EVRY (FR).

(74) Agents: CABINET ORES et al.; 6, avenue de Messine,  
F-75008 Paris (FR).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,  
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,  
VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,  
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/018820 A2

(54) Title: MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES RELATED APPLICATIONS

(57) Abstract: Processes and systems for the high throughput directed evolution of peptides and proteins, particularly those that act in complex biological settings, are provided. The proteins and peptides include, but are not limited to, intracellular proteins, messenger/signaling/hormone proteins and viral proteins. Also provided is a rational method for generating protein variants and also a method for titrating viruses.

-1-

MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES  
RELATED APPLICATIONS

Benefit of priority is claimed to U.S. application Serial No. 10/022,390, filed December 17, 2001, to Manuel Vega, Lila Drittanti and Marjorie Flaux entitled "MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES." Benefit of priority is also claimed to U.S. provisional patent application serial No. 60/315,382, filed August 27, 2001, to Manuel Vega, Lila Drittanti, and Marjorie Flaux entitled "HIGH THROUGHPUT DIRECTED EVOLUTION BY RATIONAL MUTAGENESIS." Where permitted, the subject matter of each of these applications is incorporated in its entirety by reference thereto.

**FIELD OF INVENTION**

Mutant adeno-associated virus Rep proteins, recombinant viruses that express the proteins and nucleic acid molecule encoding the Rep proteins are provided. Uses of the recombinant viruses for treatment of diseases and vectors for gene therapy are also provided.

**BACKGROUND**

Adeno-associated virus (AAV) is a defective and non-pathogenic parvovirus that requires co-infection with either adenovirus or a herpes virus, which provide helper functions, for its growth and multiplication. There is an extensive body of knowledge regarding AAV biology and genetics (see, e.g., Weitzman *et al.* (1996) *J. Virol.* 70: 2240-2248 (1996); Walker *et al.* (1997) *J. Virol.* 71:2722-2730; Urabe *et al.* (1999) *J. Virol.* 23:2682-2693; Davis *et al.* (2000) *J. Virol.* 23:74:2936-2942; Yoon *et al.* (2001) *J. Virol.* 75:3230-3239; Deng *et al.* (1992) *Anal Biochem* 200:81-85; Drittanti *et al.* (2000) *Gene Therapy* 7:924-929; Srivastava *et al.* (1983) *J. Virol.* 45:555-564; Hermonat *et al.* (1984) *J. Virol.* 51:329-339; Chejanovsky *et al.* (1989) *Virology* 173:120-128; Chejanovsky *et al.* (1990) *J. Virol.* 64:1764-1770; Owens *et al.* (1991)

-2-

- Virology* 184:14-22; Owens *et al.* (1992) *J. Virol.* 66:1236-1240; Qicheng Yang *et al.* (1992) *J. Virol.* 66:6058-6069; Qicheng Yang *et al.* (1993) *J. Virol.* 67:4442-4447; Owens *et al.* (1993) *J. Virol.* 62:997-1005; Sirkka *et al.* (1994) *J. Virol.* 68:2947-2957; Ramesh *et al.* (1995) 5 *Biochem. Biophys. Res. Com.* Vol 210 (3), 717-725; Sirkka (1995) *J. Virol.* 69:6787-6796; Sirkka *et al.* (1996) *Biochem. Biophys. Res. Com.* 220:294-299; Ryan *et al.* (1996) *J. Virol.* 70:1542-1553; Weitzman *et al.* (1996) *J. Virol.* 70:2440-2448; Walker *et al.* (1997) *J. Virol.* 71:2722-2730; Walker *et al.* (1997) *J. Virol.* 71:6996-7004; Davis *et al.* (1999) *J. Virology* 73:2084-2093; Urabe *et al.* (1999) *J. Virol.* 73:2682-2693; Gavin *et al.* (1999) *J. Virol.* 73:9433-9445; Davis *et al.* (2000) *J. Virol.* 74:2936-2942; Pei Wu *et al.* (2000) *J. Virol.* 74:8635-8647; Alessandro Marcello *et al.* (2000) *J. Virol.* 74:9090-9098). AAV are members of the family *Parvoviridae* and are assigned to the genus *Dependovirus*.
- 10 Members of this genus are small, non-enveloped, icosahedral with linear and single-stranded DNA genomes, and have been isolated from many species ranging from insects to humans.
- AAV can either remain latent after integration into host chromatin or replicate following infection. Without co-infection, AAV can enter host 15 cells and preferentially integrate at a specific site on the *q* arm of chromosome 19 in the human genome.
- The AAV genome contains 4975 nucleotides and the coding sequence is flanked by two inverted terminal repeats (ITRs) on either side that are the only sequences in *cis* required for viral assembly and 20 replication. The ITRs contain palindromic sequences, which form a hairpin secondary structure, containing the viral origins of replication. The ITRs are organized in three segments: the Rep binding site (RBS), the terminal resolution site (TRS), and a spacer region separating the RBS from 25 the TRS.

-3-

- Regulation of AAV genes is complex and involves positive and negative regulation of viral transcription. For example, the regulatory proteins Rep 78 and Rep 68 interact with viral promoters to establish a feedback loop (Beaton *et al.* (1989) *J. Virol.* 63:4450-4454; Hermonat 5 (1994) *Cancer Lett.* 81:129-136). Expression from the p5 and p19 promoters is negatively regulated in *trans* by these proteins. Rep 78 and 68, which are required for this regulation, have bind to inverted terminal repeats (ITRs; Ashktorab *et al.* (1989) *J. Virol.* 63:3034-3039) in a site- and strand-specific manner, *in vivo* and *in vitro*. This binding to ITRs 10 induces a cleavage at the TRS and permits the replication of the hairpin structure, thus, illustrating the Rep helicase and endonuclease activities (Im *et al.* (1990) *Cell* 61:447-457; and Walker *et al.* (1997) *J. Virol.* 71:6996-7004), and the role of these non-structural proteins in the initial steps of DNA replication (Hermonat *et al.* (1984) *J. Virol.* 52:329-339). 15 15 Rep 52 and 40, the two minor forms of the Rep proteins, do not bind to ITRs and are dispensable for viral DNA replication and site-specific integration (Im *et al.* (1992) *J. Virol.* 66:1119-112834; Ni *et al.* (1994) *J. Virol.* 68:1128-1138.

- The genome (see, FIG. 1) is organized into two open reading 20 frames (ORFs, designated left and right) that encode structural capsid proteins (Cap) and non-structural proteins (Rep). There are three promoters: p5 (from nucleotides 255 to 261: TATTTAA), p19 (from nucleotide 843 to 849: TATTTAA) and p40 (from nucleotides 1822 to 1827: ATATAA). The right-side ORF (see FIG. 1) encodes three capsid 25 structural proteins (Vp 1-3). These three proteins, which are encoded by overlapping DNA, result from differential splicing and the use of an unusual initiator codon (Cassineti *et al.* (1988) *Virology* 167:176-184). Expression of the capsid genes is regulated by the p40 promoter. Capsid proteins VP1, VP2 and VP3 initiate from the p40 promoter. VP1 uses an

-4-

alternate splice acceptor at nucleotide 2201; whereas VP2 and VP3 are derived from the same transcription unit, but VP2 use an ACG triplet as an initiation codon upstream from the start of VP3. On the left side of the genome, two promoters p5 and p19 direct expression of four regulatory proteins. The left flanking sequence also uses a differential splicing mechanism (Mendelson *et al.* (1986) *J. Virol.* 60:823-832) to encode the Rep proteins, designated Rep 78, 68, 52 and 40 on the basis of molecular weight. Rep 78 and 68 are translated from a transcript produced from the p5 promoter and are produced from the unspliced and spliced form, respectively, of the transcript. Rep 52 and 40 are the translation products of unspliced and spliced transcripts from the p19 promoter.

AAV and rAAV have many applications, including use as a gene transfer vector, for introducing heterologous nucleic acid into cells and for genetic therapy. Advances in the production of high-titer rAAV stocks to the transition to human clinical trials have been made, but improvement of rAAV production will be complemented with special attention to clinical applications of rAAV vectors as a successful gene therapy approach. Productivity of rAAV (i.e. the amount of vector particles that can be obtained per unitary manufacturing operation) is one of the rate limiting steps in the further development of rAAV as gene therapy vector. Methods for high throughput production and screening of rAAV have been developed (see, e.g., Drittanti *et al.* (2000) *Gene Therapy* 7:924-929). Briefly, as with the other steps in methods provided herein, the plasmid preparation, transfection, virus productivity and titer and biological activity assessment are intended to be performed in an automatable high throughput format, such as in a 96 well or loci formats (or other number of wells or multiples of 96, such as 384, 1536 . . . 9600, 9984 . . . well or loci formats).

-5-

## SUMMARY

Mutant AAV Rep proteins, nucleic acid molecules encoding such proteins, and rAAV that encode the proteins are provided. Among the rep proteins are those that result in increased rAAV production in rAAV that encode such mutants, thereby, among a variety of advantages, offering a solution to the need in the gene therapy industry to increase the production of therapeutic vectors without up-scaling manufacturing. Methods of gene therapy using the rAAV are provided.

Directed evolution methods provided in co-pending U.S. provisional application Serial No. 60/315,382, filed as U.S. application Serial No. 10/022,249, and described herein have been used to identify amino acid "hit" positions in adeno-associated virus (AAV) rep proteins that are relevant for AAV or rAAV production. Those amino acid positions are selected such that a change in the amino acid leads to a change in protein activity either to lower activity or to higher activity compared to native-sequence Rep proteins. The hit positions were then used to generate further mutants designated "leads." Provided herein are the resulting mutant rep proteins that result in either higher or lower levels of AAV or rAAV virus compared to the wild-type (native) Rep protein(s). Nucleic acid molecules that encode the mutant Rep proteins are also provided.

Also provided are rAAV that contain the nucleic acid molecules and methods that use the rAAV to produce the mutant Rep. Cell-free (*in vitro*) and intracellular methods are provided. Cells containing the rAAV are also provided.

Among the Rep mutants provided herein, in addition to Rep mutants that enhance AAV production, are those that inhibit papillomavirus (PV) and PV-associated diseases, including certain cancers and human immunodeficiency virus (HIV) and HIV-associated diseases. Methods of treating such diseases are provided.

-6-

#### DESCRIPTION OF THE FIGURES

FIGURE 1 shows the genetic map of AAV, including the location of promoters, and transcripts; amino acid 1 of the Rep 78 gene is at nucleotide 321 in the AAV-2 genome.

5 FIGURES 2A and 2B depict "HITS" and "LEADS" respectively for identification of AAV rep mutants "evolved" for increased activity.

FIGURES 3A and 3B show the alignment of amino acid sequences of Rep78 among AAV-1; AAV-6; AAV-3; AAV-3B; AAV-4; AAV-2; AAV-5 sequences, respectively; the hit positions with 100 percent homology  
10 among the serotypes are bolded italics, where the position is different (compared to AAV-2, no. 6 in the Figure) in a particular serotype, it is in bold; a sequence indicating relative conservation of sequences among the serotypes is labeled "C".

##### Legend:

15 1 is AAV-1; 2 is AAV-6, 3 is AAV-3, 4 is AAV-3B, 5 is AAV-4, 6 is AAV-2, and 7 is AAV-5;  
" ." where the amino acid is present  $\geq$  20%;  
";" where the amino acid is present  $\geq$  40%;  
"+" where the amino acid is present  $\geq$  60%;  
20 "\*" where the amino acid is present  $\geq$  80%; and where the amino acid is the same amongst all serotypes depicted it is represented by its single letter code.

#### DETAILED DESCRIPTION

25 A. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents, patent applications, published applications and publications, Genbank sequences, 30 websites and other published materials referred to throughout the entire disclosure herein are, unless noted otherwise, incorporated by reference

in their entirety. In the event that there are a plurality of definitions for terms herein, those in this section prevail.

As used herein, directed evolution refers to methods that adapt natural proteins or protein domains to work in new chemical or biological environments and/or to elicit new functions. It is more a more broad-based technology than DNA shuffling.

- As used herein, high-throughput screening (HTS) refers to processes that test a large number of samples, such as samples of test proteins or cells containing nucleic acids encoding the proteins of interest
- 10 to identify structures of interest or to identify test compounds that interact with the variant proteins or cells containing them. HTS operations are amenable to automation and are typically computerized to handle sample preparation, assay procedures and the subsequent processing of large volumes of data.
- 15 As used herein, DNA shuffling is a PCR-based technology that produces random rearrangements between two or more sequence-related genes to generate related, although different, variants of given gene.
- As used herein, "hits" are mutant proteins that have an alteration in any attribute, chemical, physical or biological property in which such
- 20 alteration is sought. In the methods herein, hits are generally generated by systematically replacing each amino acid in a protein or a domain thereof with a selected amino acid, typically Alanine, Glycine, Serine or any amino acid, as long as each residue is replaced with the same residue. Hits may be generated by other methods known to those of skill
- 25 in the art tested by the high-throughput methods herein. For purposes herein a Hit typically has activity with respect to the function of interest that differs by at least 10%, 20%, 30% or more from the wild type or native protein. The desired alteration, which is generally a reduction in activity, will depend upon the function or property of interest.

-8-

- As used herein, "leads" are "hits" whose activity has been optimized for the particular attribute, chemical, physical or biological property. In the methods herein, leads are generally produced by systematically replacing the hit loci with all remaining 18 amino acids, and
- 5    identifying those among the resulting proteins that have a desired activity.
- The leads may be further optimized by replacement of a plurality of "hit" residues. Leads may be generated by other methods known to those of skill in the art and tested by the hightthroughput methods herein. For purposes herein a lead typically has activity with respect to the function
- 10    of interest that differs from the native activity, by a desired amount and is at by at least 10%, 20%, 30% or more from the wild type or native protein. Generally a Lead will have an activity that is 2 to 10 or more times the native protein for the activity of interest. As with hits, the change in the activity is dependent upon the activity that is "evolved."
- 15    The desired alteration will depend upon the function or property of interest.

As used herein, MOI is multiplicity of infection.

As used herein, ip, with reference to a virus or recombinant vector, refers to a titer of infectious particles.

- 20    As used herein, pp refers to the total number of vector (or virus) physical particles

As used herein, biological and pharmacological activity includes any activity of a biological pharmaceutical agent and includes, but is not limited to, biological efficiency, transduction efficiency, gene/transgene expression, differential gene expression and induction activity, titer, progeny productivity, toxicity, citotoxicity, immunogenicity, cell proliferation and/or differentiation activity, anti-viral activity, morphogenetic activity, teratogenetic activity, pathogenetic activity, therapeutic activity, tumor suppressor activity, ontogenetic activity,

-9-

oncogenetic activity, enzymatic activity, pharmacological activity, cell/tissue tropism and delivery.

As used herein, "output signal" refers to parameters that can be followed over time and, if desired, quantified. For example, when a virus 5 infects or is introduced into a cell, the cell containing the virus undergoes a number of changes. Any such change that can be monitored and used to assess infection, is an output signal, and the cell is referred to as a reporter cell; the encoding nucleic acid is referred to as a reporter gene, and the construct that includes the encoding nucleic acid is a reporter 10 construct. Output signals include, but are not limited to, enzyme activity, fluorescence, luminescence, amount of product produced and other such signals. Output signals include expression of a viral gene or viral gene product, including heterologous genes (transgenes) inserted into the virus. Such expression is a function of time ("t") after infection, which in turn is 15 related to the amount of virus used to infect the cell, and, hence, the concentration of virus ("s") in the infecting composition. For higher concentrations the output signal is higher. For any particular concentration, the output signal increases as a function of time until a plateau is reached. Output signals may also measure the interaction 20 between cells, expressing heterologous genes, and biological agents

As used herein, adeno-associated virus (AAV) is a defective and non-pathogenic parvovirus that requires co-infection with either adenovirus or herpes virus for its growth and multiplication, able of providing helper functions. A variety of serotypes are known, and 25 contemplated herein. Such serotypes include, but are not limited to: AAV-1 (Genbank accession no. NC002077; accession no. VR-645); AAV-2 (Genbank accession no. NC001401; accession no. VR-680); AAV-3 (Genbank accession no. NC001729; accession no. VR-681); AAV-3b (Genbank accession no. NC001863); AAV-4 (Genbank accession no.

-10-

NC001829; ATCC accession no. VR-646 ); AAV-6 (Genbank accession no.NC001862); and avian associated adeno-virus (ATCC accession no. VR-1449). The preparation and use of AAVs as vectors for gene expression *in vitro* and for *in vivo* use for gene therapy are well known 5 (see, e.g., U.S. Patent Nos. 4,797,368, 5,139,941, 5,798,390 and 6,127,175; Tessier *et al.* (2001) *J. Virol.* 75:375-383; Salvetti *et al.* (1998) *Hum Gene Ther* 20:695-706; Chadeuf *et al.* (2000) *J Gene Med* 2:260-268).

As used herein, the activity of a Rep protein or of a capsid protein 10 refers to any biological activity that can be assessed. In particular, herein, the activity assessed for the rep proteins is the amount (*i.e.*, titer) of AAV produced by a cell.

As used herein, the Hill equation is a mathematical model that relates the concentration of a drug (*i.e.*, test compound or substance) to 15 the response being measured

$$y = \frac{y_{\max}[D]^n}{[D]^n + [D_{50}]^n},$$

where  $y$  is the variable being measured, such as a response, signal,  $y_{\max}$  is 20 the maximal response achievable,  $[D]$  is the molar concentration of a drug,  $[D_{50}]$  is the concentration that produces a 50% maximal response to the drug,  $n$  is the slope parameter, which is 1 if the drug binds to a single 25 site and with no cooperativity between or among sites. A Hill plot is  $\log_{10}$  of the ratio of ligand-occupied receptor to free receptor vs.  $\log [D]$  (M). The slope is  $n$ , where a slope of greater than 1 indicates cooperativity among binding sites, and a slope of less than 1 can indicate heterogeneity of binding. This general equation has been employed for assessing 30 interactions in complex biological systems (see, published International

-11-

PCT application No. WO 01/44809 based on PCT n° PCT/FR00/03503, see, also, EXAMPLES).

As used herein, in the Hill-based analysis (published International PCT application No. WO 01/44809 based on PCT n° PCT/FR00/03503),

5 the parameters,  $\pi, \kappa, \tau, \epsilon, \eta, \theta$ , are as follows:

$\pi$  potency of the biological agent acting on the assay (cell-based) system;

$\kappa$  constant of resistance of the assay system to elicit a response to a biological agent;

10  $\epsilon$  is global efficiency of the process or reaction triggered by the biological agent on the assay system;

$\tau$  is the apparent titer of the biological agent;

$\theta$  is the absolute titer of the biological agent; and

$\eta$  is the heterogeneity of the biological process or reaction.

15 In particular, as used herein, the parameters  $\pi$  (potency) or  $\kappa$  (constant of resistance) are used to respectively assess the potency of a test agent to produce a response in an assay system and the resistance of the assay system to respond to the agent.

As used herein,  $\epsilon$  (efficiency), is the slope at the inflection point of  
20 the Hill curve (or, in general, of any other sigmoidal or linear approximation), to assess the efficiency of the global reaction (the biological agent and the assay system taken together) to elicit the biological or pharmacological response.

As used herein,  $\tau$  (apparent titer) is used to measure the limiting dilution or the apparent titer of the biological agent.

As used herein,  $\theta$  (absolute titer), is used to measure the absolute limiting dilution or titer of the biological agent.

As used herein,  $\eta$  (heterogeneity) measures the existence of discontinuous phases along the global reaction, which is reflected by an

-12-

abrupt change in the value of the Hill coefficient or in the constant of resistance.

As used herein, a library of mutants refers to a collection of plasmids or other vehicles that carry (encode) the gene variants, such that

5 individual plasmids or other vehicles carry individual gene variants.

When a library of proteins is contemplated, it will be so-stated.

As used herein, a "reporter cell" is the cell that "reports", *i.e.*, undergoes the change, in response to introduction of the nucleic acid infection and, therefore, it is named here a reporter cell.

10 As used herein, "reporter" or "reporter moiety" refers to any moiety that allows for the detection of a molecule of interest, such as a protein expressed by a cell. Reporter moieties include, but are not limited to, for example, fluorescent proteins, such as red, blue and green fluorescent proteins; lacZ and other detectable proteins and gene products. For  
15 expression in cells, nucleic acid encoding the reporter moiety can be expressed as a fusion protein with a protein of interest or under the control of a promoter of interest.

As used herein, a titering virus increases or decreases the output signal from a reporter virus, which is a virus that can be detected, such  
20 as by a detectable label or signal.

As used herein, phenotype refers to the physical, physiological or other manifestation of a genotype (a sequence of a gene). In methods herein, phenotypes that result from alteration of a genotype are assessed.

As used herein, activity refers to the function or property to be  
25 evolved. An active site refers to a site(s) responsible or that participates in conferring the activity or function. The activity or active site evolved (the function or property and the site conferring or participating in conferring the activity) may have nothing to do with natural activities of

-13-

a protein. For example, it could be an 'active site' for conferring immunogenicity (immunogenic sites or epitopes) on a protein.

As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their known,

- 5 three-letter or one-letter abbreviations (see, Table 1). The nucleotides, which occur in the various nucleic acid fragments, are designated with the standard single-letter designations used routinely in the art.

As used herein, amino acid residue refers to an amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide

- 10 linkages. The amino acid residues described herein are presumed to be in the "L" isomeric form. Residues in the "D" isomeric form, which are so-designated, can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH<sub>2</sub> refers to the free amino group present at the amino terminus of a polypeptide.
- 15 COOH refers to the free carboxy group present at the carboxyl terminus of a polypeptide. In keeping with standard polypeptide nomenclature described in *J. Biol. Chem.*, 243:3552-59 (1969) and adopted at 37 C.F.R. §§ 1.821 - 1.822, abbreviations for amino acid residues are shown in the following Table:

20

**Table 1**  
**Table of Correspondence**

25

SYMBOL		
1-Letter	3-Letter	AMINO ACID
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine

-14-

SYMBOL		
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
X	Xaa	Unknown or other

- It should be noted that all amino acid residue sequences  
 20 represented herein by formulae have a left to right orientation in the conventional direction of amino-terminus to carboxyl-terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those referred to in 37 C.F.R. §§ 1.821-  
 25 1.822, and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more

-15-

amino acid residues or to an amino-terminal group such as NH<sub>2</sub> or to a carboxyl-terminal group such as COOH.

In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally

5 without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. *Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. co., p.224).

10 Such substitutions are preferably made in accordance with those set forth in TABLE 2 as follows:

**TABLE 2**

	Original residue	Conservative substitution
15	Ala (A)	Gly; Ser
	Arg (R)	Lys
	Asn (N)	Gln; His
	Cys (C)	Ser
	Gln (Q)	Asn
20	Glu (E)	Asp
	Gly (G)	Ala; Pro
	His (H)	Asn; Gln
	Ile (I)	Leu; Val
	Leu (L)	Ile; Val
25	Lys (K)	Arg; Gln; Glu
	Met (M)	Leu; Tyr; Ile
	Phe (F)	Met; Leu; Tyr
	Ser (S)	Thr
	Thr (T)	Ser
30	Trp (W)	Tyr
	Tyr (Y)	Trp; Phe
	Val (V)	Ile; Leu

Other substitutions are also permissible and may be determined empirically or in accord with known conservative substitutions.

As used herein, nucleic acids include DNA, RNA and analogs

35 thereof, including protein nucleic acids (PNA) and mixture thereof.

Nucleic acids can be single or double stranded. When referring to probes

-16-

or primers, optionally labeled, with a detectable label, such as a fluorescent or radiolabel, single-stranded molecules are contemplated. Such molecules are typically of a length such that they are statistically unique of low copy number (typically less than 5, preferably less than 3)

- 5 for probing or priming a library. Generally a probe or primer contains at least 14, 16 or 30 contiguous of sequence complementary to or identical to a gene of interest. Probes and primers can be 10, 14, 16, 20, 30, 50, 100 or more nucleic acid bases long.

- As used herein, homologous means about greater than 25% nucleic acid sequence identity, preferably 25% 40%, 60%, 80%, 90% or 95%.  
10 The intended percentage will be specified. The terms "homology" and "identity" are often used interchangeably. In general, sequences are aligned so that the highest order match is obtained (see, e.g.:  
*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University  
15 Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*,  
20 Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; Carillo *et al.* (1988) *SIAM J Applied Math* 48:1073). By sequence identity, the number of conserved amino acids are determined by standard alignment algorithms programs, and are used with default gap penalties established by each supplier. Substantially homologous nucleic acid  
25 molecules would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid of interest. Also contemplated are nucleic acid molecules that contain degenerate codons in place of codons in the hybridizing nucleic acid molecule.

-17-

As used herein, a nucleic acid homolog refers to a nucleic acid that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" it is meant having at least 80%, preferably at least 90%,  
5 most preferably at least 95% homology therewith or a less percentage of homology or identity and conserved biological activity or function.

The terms "homology" and "identity" are often used interchangeably. In this regard, percent homology or identity may be determined, for example, by comparing sequence information using a GAP  
10 computer program. The GAP program uses the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443 (1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482 (1981). Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total number of symbols in.  
15 the shorter of the two sequences. The preferred default parameters for the GAP program may include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745 (1986), as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN*  
20 *SEQUENCE AND STRUCTURE*, National Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

Whether any two nucleic acid molecules have nucleotide sequences  
25 that are, for example, at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% "identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988).

-18-

Alternatively the BLAST function of the National Center for Biotechnology Information database may be used to determine identity

- In general, sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heijne, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988)). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988). Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(II):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., et al., *J Molec Biol* 215:403 (1990)), and CLUSTALW. For sequences displaying a relatively high degree of homology, alignment can be effected manually by simply lining up the sequences by eye and matching the conserved portions.

-19-

Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide. For example, a test polypeptide may be defined as any polypeptide that is 90% or more identical to a reference polypeptide.

- 5 For the alignments presented herein (see, Figures 3A and 3B) for the AAV serotype, the CLUSTALW program was employed with parameters set as follows: scoring matrix BLOSUM, gap open 10, gap extend 0.1, gap distance 40% and transitions/transversions 0.5; specific residue penalties for hydrophobic amino acids (DEGKNPQRS), distance  
10 between gaps for which the penalties are augmented was 8, and gaps of extemeties penalized less than internal gaps.

- As used herein, a "corresponding" position on a protein, such as the AAV rep protein, refers to an amino acid position based upon alignment to maximize sequence identity. For AAV Rep proteins an  
15 alignment of the Rep 78 protein from AAV-2 and the corresponding protein from other AAV serotypes (AAV-1, AAV-6, AAV-3, AAV-3B, AAV-4, AAV-2 and AAV-5) is shown in Figures 3A and 3B. The "hit" positions are shown in italics.

- As used herein, the term at least "90% identical to" refers to  
20 percent identities from 90 to 100% relative to the reference polypeptides. Identity at a level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared. No more than 10% (i.e., 10 out of 100) amino acids in the test polypeptide differs from that of the reference  
25 polypeptides. Similar comparisons may be made between a test and reference polynucleotides. Such differences may be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they may be clustered in one or more locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference

-20-

(approximately 90% identity). Differences are defined as nucleic acid or amino acid substitutions, or deletions.

As used herein, it is also understood that the terms substantially identical or similar varies with the context as understood by those skilled

5 in the relevant art.

As used herein, genetic therapy involves the transfer of heterologous nucleic acids to the certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The nucleic acid, such as DNA, is introduced into the selected

- 10 target cells in a manner such that the heterologous nucleic acid, such as DNA, is expressed and a therapeutic product encoded thereby is produced. Alternatively, the heterologous nucleic acid, such as DNA, may in some manner mediate expression of DNA that encodes the therapeutic product, or it may encode a product, such as a peptide or
- 15 RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic acid encoding a gene product that replaces a defective gene or supplements a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic
- 20 compound, such as a growth factor or inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous nucleic acid, such as DNA, encoding the therapeutic product
- 25 may be modified prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof. Genetic therapy may also involve delivery of an inhibitor or repressor or other modulator of gene expression.

-21-

- As used herein, heterologous or foreign nucleic acid, such as DNA and RNA, are used interchangeably and refer to DNA or RNA that does not occur naturally as part of the genome in which it is present or which is found in a location or locations in the genome that differ from that in which it occurs in nature. Heterologous nucleic acid is generally not endogenous to the cell into which it is introduced, but has been obtained from another cell or prepared synthetically. Generally, although not necessarily, such nucleic acid encodes RNA and proteins that are not normally produced by the cell in which it is expressed. Any DNA or RNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by heterologous DNA. Heterologous DNA and RNA may also encode RNA or proteins that mediate or alter expression of endogenous DNA by affecting transcription, translation, or other regulatable biochemical processes.
- Examples of heterologous nucleic acid include, but are not limited to, nucleic acid that encodes traceable marker proteins, such as a protein that confers drug resistance, nucleic acid that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and DNA that encodes other types of proteins, such as antibodies.
- Hence, herein heterologous DNA or foreign DNA, includes a DNA molecule not present in the exact orientation and position as the counterpart DNA molecule found in the genome. It may also refer to a DNA molecule from another organism or species (*i.e.*, exogenous).
- As used herein, a therapeutically effective product introduced by genetic therapy is a product that is encoded by heterologous nucleic acid, typically DNA, that, upon introduction of the nucleic acid into a host, a product is expressed that ameliorates or eliminates the symptoms, manifestations of an inherited or acquired disease or that cures the disease.

-22-

As used herein, a therapeutically effective dose refers to that amount of the compound sufficient to result in amelioration of symptoms of disease.

As used herein, isolated with reference to a nucleic acid molecule or polypeptide or other biomolecule means that the nucleic acid or polypeptide has separated from the genetic environment from which the polypeptide or nucleic acid were obtained. It may also mean altered from the natural state. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Thus, a polypeptide or polynucleotide produced and/or contained within a recombinant host cell is considered isolated. Also intended as an "isolated polypeptide" or an "isolated polynucleotide" are polypeptides or polynucleotides that have been purified, partially or substantially, from a recombinant host cell or from a native source. For example, a recombinantly produced version of a compound can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). The terms isolated and purified are sometimes used interchangeably.

Thus, by "isolated" it is meant that the nucleic acid is free of the coding sequences of those genes that, in the naturally-occurring genome of the organism (if any), immediately flank the gene encoding the nucleic acid of interest. Isolated DNA may be single-stranded or double-stranded, and may be genomic DNA, cDNA, recombinant hybrid DNA, or synthetic DNA. It may be identical to a native DNA sequence, or may differ from such sequence by the deletion, addition, or substitution of one or more nucleotides.

Isolated or purified as it refers to preparations made from biological cells or hosts means any cell extract containing the indicated DNA or

-23-

protein including a crude extract of the DNA or protein of interest. For example, in the case of a protein, a purified preparation can be obtained following an individual technique or a series of preparative or biochemical techniques and the DNA or protein of interest can be present at various 5 degrees of purity in these preparations. The procedures may include for example, but are not limited to, ammonium sulfate fractionation, gel filtration, ion exchange change chromatography, affinity chromatography, density gradient centrifugation and electrophoresis.

A preparation of DNA or protein that is "substantially pure" or  
10 "isolated" should be understood to mean a preparation free from naturally occurring materials with which such DNA or protein is normally associated in nature. "Essentially pure" should be understood to mean a "highly" purified preparation that contains at least 95% of the DNA or protein of interest.

15 A cell extract that contains the DNA or protein of interest should be understood to mean a homogenate preparation or cell-free preparation obtained from cells that express the protein or contain the DNA of interest. The term "cell extract" is intended to include culture media, especially spent culture media from which the cells have been removed.

20 As used herein, receptor refers to a biologically active molecule that specifically binds to (or with) other molecules. The term "receptor protein" may be used to more specifically indicate the proteinaceous nature of a specific receptor.

As used herein, recombinant refers to any progeny formed as the  
25 result of genetic engineering.

As used herein, a promoter region refers to the portion of DNA of a gene that controls transcription of the DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for RNA polymerase recognition, binding and transcription

-24-

initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, binding and transcription initiation activity of the RNA polymerase. These sequences may be *cis* acting or may be

- 5 responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated.

As used herein, the phrase "operatively linked" generally means the sequences or segments have been covalently joined into one piece of DNA, whether in single or double stranded form, whereby control or

- 10 regulatory sequences on one segment control or permit expression or replication or other such control of other segments. The two segments are not necessarily contiguous. For gene expression a DNA sequence and a regulatory sequence(s) are connected in such a way to control or permit gene expression when the appropriate molecules, e.g., transcriptional activator proteins, are bound to the regulatory sequence(s).

As used herein, production by recombinant means by using recombinant DNA methods means the use of the well known methods of molecular biology for expressing proteins encoded by cloned DNA, including cloning expression of genes and methods, such as gene

- 15 shuffling and phage display with screening for desired specificities.

As used herein, a splice variant refers to a variant produced by differential processing of a primary transcript of genomic DNA that results in more than one type of mRNA.

- As used herein, a composition refers to any mixture of two or more products or compounds. It may be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

-25-

As used herein, a combination refers to any association between two or more items.

As used herein, substantially identical to a product means sufficiently similar so that the property of interest is sufficiently

- 5 unchanged so that the substantially identical product can be used in place of the product.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked.

- One type of preferred vector is an episome, i.e., a nucleic acid capable of
- 10 extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques
- 15 are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form, are not bound to the chromosome. "Plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. Other such other forms of expression vectors that serve equivalent functions and that
- 20 become known in the art can be used subsequently hereto.

As used herein, vector is also used interchangeable with "virus vector" or "viral vector". In this case, which will be clear from the context, the "vector" is not self-replicating. Viral vectors are engineered viruses that are operatively linked to exogenous genes to transfer (as

- 25 vehicles or shuttles) the exogenous genes into cells.

As used herein, transduction refers to the process of gene transfer and expression into mammalian and other cells mediated by viruses.

Transfection refers to the process when mediated by plasmids.

-26-

As used herein, "polymorphism" refers to the coexistence of more than one form of a gene or portion thereof. A portion of a gene of which there are at least two different forms, i.e., two different nucleotide sequences, is referred to as a "polymorphic region of a gene". A

- 5 polymorphic region can be a single nucleotide, referred to as a single nucleotide polymorphism (SNP), the identity of which differs in different alleles. A polymorphic region can also be several nucleotides in length.

As used herein, "polymorphic gene" refers to a gene having at least one polymorphic region.

- 10 As used herein, "allele", which is used interchangeably herein with "allelic variant" refers to alternative forms of a gene or portions thereof.

Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be homozygous for the gene or allele. When a subject has two different

- 15 alleles of a gene, the subject is said to be heterozygous for the gene.

Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing a mutation.

- 20 As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid molecule comprising an open reading frame and including at least one exon and (optionally) an intron sequence. A gene can be either RNA or DNA. Genes may include regions preceding and following the coding region (leader and trailer).

- 25 As used herein, "intron" refers to a DNA sequence present in a given gene which is spliced out during mRNA maturation.

As used herein, "nucleotide sequence complementary to the nucleotide sequence set forth in SEQ ID NO: x" refers to the nucleotide sequence of the complementary strand of a nucleic acid strand having

-27-

SEQ ID NO: x. The term "complementary strand" is used herein interchangeably with the term "complement". The complement of a nucleic acid strand can be the complement of a coding strand or the complement of a non-coding strand. When referring to double stranded 5 nucleic acids, the complement of a nucleic acid having SEQ ID NO: x refers to the complementary strand of the strand having SEQ ID NO: x or to any nucleic acid having the nucleotide sequence of the complementary strand of SEQ ID NO: x. When referring to a single stranded nucleic acid having the nucleotide sequence SEQ ID NO: x, the complement of this 10 nucleic acid is a nucleic acid having a nucleotide sequence which is complementary to that of SEQ ID NO: x.

As used herein, the term "coding sequence" refers to that portion of a gene that encodes an amino acid sequence of a protein.

As used herein, the term "sense strand" refers to that strand of a 15 double-stranded nucleic acid molecule that has the sequence of the mRNA that encodes the amino acid sequence encoded by the double-stranded nucleic acid molecule.

As used herein, the term "antisense strand" refers to that strand of a double-stranded nucleic acid molecule that is the complement of the 20 sequence of the mRNA that encodes the amino acid sequence encoded by the double-stranded nucleic acid molecule.

As used herein, an array refers to a collection of elements, such as nucleic acid molecules, containing three or more members. An addressable array is one in which the members of the array are 25 identifiable, typically by position on a solid phase support or by virtue of an identifiable or detectable label, such as by color, fluorescence, electronic signal (*i.e.* RF, microwave or other frequency that does not substantially alter the interaction of the molecules of interest), bar code or other symbology, chemical or other such label. Hence, in general the

-28-

members of the array are immobilized to discrete identifiable loci on the surface of a solid phase or directly or indirectly linked to or otherwise associated with the identifiable label, such as affixed to a microsphere or other particulate support (herein referred to as beads) and suspended in  
5 solution or spread out on a surface.

As used herein, a support (also referred to as a matrix support, a matrix, an insoluble support or solid support) refers to any solid or semisolid or insoluble support to which a molecule of interest, typically a biological molecule, organic molecule or biospecific ligand is linked or  
10 contacted. Such materials include any materials that are used as affinity matrices or supports for chemical and biological molecule syntheses and analyses, such as, but are not limited to: polystyrene, polycarbonate, polypropylene, nylon, glass, dextran, chitin, sand, pumice, agarose, polysaccharides, dendrimers, buckyballs, polyacrylamide, silicon, rubber,  
15 and other materials used as supports for solid phase syntheses, affinity separations and purifications, hybridization reactions, immunoassays and other such applications. The matrix herein can be particulate or can be in the form of a continuous surface, such as a microtiter dish or well, a glass slide, a silicon chip, a nitrocellulose sheet, nylon mesh, or other  
20 such materials. When particulate, typically the particles have at least one dimension in the 5-10 mm range or smaller. Such particles, referred collectively herein as "beads", are often, but not necessarily, spherical. Such reference, however, does not constrain the geometry of the matrix, which may be any shape, including random shapes, needles, fibers, and  
25 elongated. Roughly spherical "beads", particularly microspheres that can be used in the liquid phase, are also contemplated. The "beads" may include additional components, such as magnetic or paramagnetic particles (see, e.g., Dyna beads (Dynal, Oslo, Norway)) for separation

-29-

using magnets, as long as the additional components do not interfere with the methods and analyses herein.

As used herein, matrix or support particles refers to matrix materials that are in the form of discrete particles. The particles have any

5 shape and dimensions, but typically have at least one dimension that is 100 mm or less, 50 mm or less, 10 mm or less, 1 mm or less, 100  $\mu\text{m}$  or less, 50  $\mu\text{m}$  or less and typically have a size that is 100  $\text{mm}^3$  or less, 50  $\text{mm}^3$  or less, 10  $\text{mm}^3$  or less, and 1  $\text{mm}^3$  or less, 100  $\mu\text{m}^3$  or less and may be order of cubic microns. Such particles are collectively called "beads."

10 As used herein, the abbreviations for any protective groups, amino acids and other compounds, are, unless indicated otherwise, in accord with their common usage, recognized abbreviations, or the IUPAC-IUB Commission on Biochemical Nomenclature (see, (1972) *Biochem.* 11:942-944).

15 **B. DIRECTED EVOLUTION OF A VIRAL GENE**

Recombinant viruses have been developed for use as gene therapy vectors. Gene therapy applications are hampered by the need for development of vectors with traits optimized for this application. The high throughput methods provided herein are ideally suited for

20 development of such vectors. In addition to use for development of recombinant viral vectors for gene therapy, these methods can also be used to study and modify the viral vector backbone architecture, trans-complementing helper functions, where appropriate, regulatable and tissue specific promoters and transgene and genomic sequence analyses.

25 Recombinant AAV (rAAV) is a gene therapy vector that can serve these and other purposes.

The rep protein is an adeno-associated virus protein involved in a number of biological processes necessary to AAV replication. The production of the rRep proteins enables viral DNA to replicate,

-30-

encapsulate and integrate (McCarty *et al.* (1992) *J. Virol* 66:4050-4057; Horer *et al.* (1995) *J. Virol* 69:5485-5496, Berns *et al.* (1996) Biology of Adeno-associated virus, in Adeno-associated virus (AAV) Vectors in Gene Therapy, K.I. Berns and C. Giraud, Springer (1996); and Chiorini *et al.*

5     (1996) The Roles of AAV Rep Proteins in gene Expression and Targeted Integration, *from* Adeno-associated virus (AAV) Vectors in Gene Therapy, K.I. Berns and C. Giraud, Springer (1996)). A rep protein with improved activity could lead to increased amounts of virus progeny thus allowing higher productivity of rAAV vectors.

10    Since the Rep protein is involved in replication it can serve as a target for increasing viral production. Since it has a variety of functions and its role in replication is complex, it has heretofore been difficult to identify mutations that result in increase viral production. The methods herein, which rely on *in vivo* screening methods, permit optimization of its 15 activities as assessed by increases in viral production. Provided herein are Rep proteins and viruses and viral vectors containing the mutated Rep proteins that provide such increase. The amino acid positions on the rep proteins that are relevant for rep proteins activities in terms of AAV or rAAV virus production are provided. Those amino acid positions are such 20 that a change in the amino acid leads to a change in protein activity either to lower activity or increase activity. As shown herein, the alanine or amino acid scan revealed the amino acid positions important for such activity (i.e. hits). Subsequent mutations produced by systematically replacing the amino acids at the hit positions with the remaining 18 amino 25 acids produced so-called "leads" that have amino acid changes and result in higher virus production. In this particular example, the method used included the following specific steps.

**Amino acid scan**

In order to first identify those amino acid (aa) positions on the rep protein that are involved in rep protein activity, an Ala-scan was performed on the rep sequence. For this, each aa in the rep protein sequence was individually changed to Alanine. Any other amino acid,

- 5 particularly another amino acid such as Gly or Ser that has a neutral effect on structure, could have been used. Each resulting mutant rep protein was then expressed and the amount of virus it produced was measured. The relative activity of each individual mutant compared to the native protein is indicated in FIG 2A. HITS are those mutants that
- 10 produce a decrease in the activity of the protein (in the example: all the mutants with activities below about 20 % of the native activity).

In a second experimental round, which included a new set of mutations and phenotypic analysis, each amino acid position hit by the Ala-scan step, was mutated by amino acid replacement of the native

- 15 amino acid by the remaining 18 amino acids, using site-directed mutagenesis.

- In both rounds, each mutant was individually designed, generated and processed separately, and optionally in parallel with the other mutants. Neither combinatorial generation of mutants nor mixtures
- 20 thereof were used in any step of the method.

A plasmid library was thus generated in which each plasmid contained a different mutant bearing a different amino acid at a different hit position. Again, each resulting mutant rep protein was then expressed and the amount of virus it could produce measured as indicated below.

- 25 The relative activity of each individual mutant compared to the native protein is indicated in FIGURE 2B. LEADS are those mutants that lead to an increase in the activity of the protein (in the example: the ten mutants with activities higher, typically between 2 to 10 times or more, generally 6-10 time, than the native activity).

-32-

**Expression of the genetic variants and phenotypic characterization.**

The rep protein acts as an intracellular protein through complex interaction with a molecular network composed by cellular proteins, DNA, AAV proteins and adenoviral proteins (note: some adenovirus proteins have to be present for the rep protein to work). The final outcome of the rep protein activity is the virus offspring composed by infectious rAAV particles. It can be expected that the activity of rep mutants would affect the titer of the rAAV virus coming out of the cells.

As the phenotypic characterization of the rep variants can only be accomplished by assaying its activity from inside mammalian cells, a mammalian cell-based expression system as well as a mammalian cell-based assay was used. The individual rep protein variants were expressed in human 293 HEK cells, by transfection of the individual plasmids constituting the diverse plasmid library. All necessary functions were provided as follows:

- (a) the cellular proteins present in the permissive specific 293 HEK cells;
- (b) the AAV necessary proteins and DNA were provided by co-transfection of the AAV cap gene as well as a rAAV plasmid vector
- 20 providing the necessary signaling and substrate ITRs sequences;
- (c) the adenovirus (AV) proteins were provided by co-transfection with a plasmid expressing all the AV helper functions.

A library of recombinant viruses with mutant rep encoding genes was generated. Each recombinant, upon introduction into a mammalian cell and expression resulted in production of rAAV infectious particles. The number of infectious particles produced by each recombinant was determined in order to assess the activity of the rep variant that had generated that amount of infectious particles.

-33-

The number of infectious particles produced was determined in a cell-based assay in which the activity of a reporter gene, in the exemplified embodiment, the bacterial lacZ gene, or virus replication (Real time PCR) was performed to quantitatively assess the number of viruses.

- 5     The limiting dilution (titer) for each virus preparation (each coming from a different rep variant) was determined by serial dilution of the viruses produced, followed by infection of appropriate cells (293 HEK or HeLa rep/cap 32 cells) with each dilution for each virus and then by measurement of the activity of the reporter gene for each dilution of each
- 10    virus. Hill plots (NAUTSCAN") (published as International PCT application No. WO 01/44809 based on PCT n° PCT/FR00/03503, Dec, 2000; see EXAMPLES) or a second order polynomial function (Drittanti *et al.* (2000) *Gene Ther.* 7: 924-929; see co-pending U.S. provisional application Serial 60/315,382) was used to analyze the readout data and to calculate the
- 15    virus titers. Briefly, the titer was calculated from the second order polynomial function by non-linear regression fitting of the experimental data. The point where the polynomial curve reaches its minimum is considered to be the titer of the rAAV preparation. Results are shown in the EXAMPLE below.
- 20    **Comparison between results of full-length Hit position analysis reporter here and the literature**  
The experiments identified a number of heretofore unknown mutation loci, which include the hits at positions: 4, 20, 22, 28, 32, 38, 39, 54, 59, 124, 125, 127, 132, 140, 161, 163, 193, 196, 197, 221,
- 25    228, 231, 234, 258, 260, 263, 264, 334, 335, 341, 342, 347, 350, 354, 363, 364, 367, 370, 376, 381, 389, 407, 411, 414, 420, 421, 422, 428, 429, 438, 440, 451, 460, 462, 484, 488, 495, 497, 498, 499, 503, 511, 512, 516, 517 and 518 with reference to the amino acids in Rep78 and Rep 68. Rep 78 is encoded by nucleotides 321-

-34-

2,186; Rep 68 is encoded by nucleotides 321-1906 and 2228-2252; Rep 52 is encoded by nucleotides 993-2186, and Rep 40 is encoded by amino acids 993-1906 and 2228-2252 of wildtype AAV.

Also among these are mutations that may have multiple effects.

- 5 Since the Rep coding region is quite complex, some of the mutations have several effects. Amino acids 542, 598, 600 and 601, which are in the to the Rep 68 and 40 intron region, are also in the coding region of Rep 78 and 52. Codon 630 is in the coding region of Rep 68 and 40 and non coding region of Rep 78 and 52.
- 10 Mutations at 10, 86, 101, 334 and 519 have been previously identified, and mutations, at loci 64, 74, 88, 175, 237, 250 and 429, but with different amino acid substitutions, have been previously reported. In all instances, however, the known mutations reportedly decrease the activity of Rep proteins. Among mutations described herein, are
- 15 mutations that result in increases in the activity the Rep function as assessed by detecting increased AAV production.

In particular, as described in the Example, mutations in the Rep-encoding region of AAV, including serotypes AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6 are provided (see Example below).

- 20 The mutant proteins and mutant adeno-associate virus (AAV) Rep proteins are provided. Exemplary proteins with mutations at one or more of residues 4, 20, 22, 29, 32, 38, 39, 54, 59, 124, 125, 127, 132, 140, 161, 163, 193, 196, 197, 221, 228, 231, 234, 258, 260, 263, 264, 334, 335, 337, 342, 347, 350, 354, 363, 364, 367, 370, 376, 381,
- 25 389, 407, 411, 414, 420, 421, 422, 424, 428, 438, 440, 451, 460, 462, 484, 488, 495, 497, 498, 499, 503, 511, 512, 516, 517, 518, 542, 548, 598, 600 and 601 of AAV-2 or the corresponding residues in other serotypes are provided. Residue 1 corresponds to residue 1 of the Rep78 protein encoded by nucleotides 321-323 of the AAV-2 genome

(see Figure 3 and the Table below for an alignment of the mutations from various serotypes).

Of particular interest are mutations that increase activity of the Rep proteins compared to wildtype. Such mutations include one or more of residues 350, 462, 497, 517, 542, 548, 598, 600 and 630 of AAV-2 and the corresponding residues in other serotypes. Also provided are mutations at or near those residues, such as within about 1 to about 10 residues of these residues such that the resulting protein has increased activity. Mutations include insertions, deletions and replacements.

**10 Lead identification.**

Based on the results obtained from the assays described herein (i.e. titer of virus produced by each rep variant), each individual rep variant was assigned a specific activity. Those variant proteins displaying the highest titers were selected as leads and are used to produce rAAV.

**15** In further steps, rAAV and Rep proteins that contain a plurality of mutations based on the hits (see Table in the EXAMPLE, listing the hits and lead sites), are produced to produce rAAV and Rep proteins that have activity that is further optimized. Examples of such proteins and AAV containing such proteins are described in the EXAMPLE. Other

**20** combinations of mutations can be prepared and tested as described herein to identify other leads of interest, particularly those that have increased Rep protein activity or that result in higher viral titers in cells containing such viruses that include appropriate *cis* acting elements for viral production.

**25** The rAAV rep mutants are used as expression vectors, which, for example, can be used transiently for the production of recombinant AAV stocks. Alternatively, the recombinant plasmids may be used to generate stable packaging cell lines.

-36-

Also among the uses of rAAV, particularly the high titer stocks produced herein, is gene therapy for the purpose of transferring genetic information into appropriate host cells for the management and correction of human diseases including inherited and acquired disorders such as

- 5 cancer and AIDS. The rAAV can be administered to a patient at therapeutically effective doses.

**C. Uses of the mutant Rep genes and the rAAV**

**Gene therapy**

- The rAAV provided herein are intended for use as vectors for gene therapy. The rAAV provided herein are intended for use in any gene therapy protocol that uses AAV as a vector. The mutant Rep proteins and nucleic acid molecules can be used to replace the corresponding gene in other AAV vectors. Of interest are the mutations provided herein that increase rAAV production. In particular, the mutant Rep proteins are used
- 10 to increase production of rAAV derived from any of the AAV serotypes, including AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6 serotypes.

- Toxicity and therapeutic efficacy of the rAAV can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LDS<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD<sub>50</sub>/ED<sub>50</sub>. Doses that exhibit large therapeutic indices are preferred. Doses that
- 20 exhibit toxic side effects may be used; care should be taken to design a delivery system that targets rAAV to the site of treatment in order to minimize damage to untreated cells and reduce side effects.

The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage

-37-

- of such rAAV lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. A therapeutically effective dose can be
- 5 estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC<sub>50</sub> (ie., the concentration of the test compound which achieves a half-maximal infection or a half- maximal inhibition) as determined in cell culture. Such information can be used to more
- 10 accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

**Treatment of Cancer, HIV, and papilloma and herpes virus infections and diseases mediated thereby**

- AAV, which is a helper-dependent parvovirus requires co-infection
- 15 with an adenovirus, herpes virus or papilloma virus (PV) for replication and particle formation. AAV inhibits PV-induced oncogenic transformation, and this inhibition has been mapped to the Rep78 protein. The Rep78 protein inhibits expression of the PV promoter just upstream of the E6 gene (p89 of bovine PV-1 (BPV-1)) p97 of human PV-16 (HPV-20), and p105 of human PV-18 (HPV-18)). DNA binding is required for this inhibition. Rep78 also binds to the TAR sequences (nt +23 to +42) and to a region just upstream of the TATA box (nt. -54 to -34) in the HIV LTR region. AAV Rep78 also regulates a variety of other cancer associated genes, including, but are not limited to, C-H-ras (Khleif *et al.*
- 25 (1991) *Virology* 181:738-741), c-fos and c-myc (Hermonat (1994) *Cancer Letts* 81:129-136).

Infection by AAV is negatively associated with cervical cancer. Infection and DNA integration by certain PV types are central events in the etiology of cervical cancer (Durst *et al.* (1983) *Proc. Natl. Acad. Sci.*

*U.S.A. 80:3812-3815; Cullen et al. (1991) J. Virol. 65:606-612).*

Roughly two thirds of cervical cancers contain the HPV-16 virus. AAV is also commonly found in the anogenital region (Han et al. (1996) *Virus Genes* 12:47-52.

- 5 Contemplated herein are AAV rep mutants that bind with greater affinity than wild-type AAV Rep78 to nucleic acid from PV, AAV, oncogenes or HIV, particularly HIV-1, and particularly promoter and other transcriptional/translational regulatory sequences from these sources. The mutant Rep protein when administered to a subject can inhibit PV
- 10 and PV-associated diseases, HIV and HIV-associated diseases. Hence methods for treatment of PV and HIV-mediated disorders by administration of rAAV encoding mutant the Rep78 genes are provided. The particular mutants for use in these methods can be identified by testing each mutant for inhibitory activity, for example, in cell-based
- 15 assays. For example, the Rep mutant protein can be tested by contacting it with nucleic acid from a PV, AAV or HIV or oncogene for a time sufficient to permit binding thereto, and comparing such binding to the binding of a wild-type Rep protein under the same conditions. Alternatively competitive binding assays may be performed. Mutant
- 20 proteins having higher binding affinities are identified.

Fusion proteins containing a *tat* protein of HIV or other targeting agent and mutant Rep protein are also provided. Pharmaceutical compositions containing such fusion proteins are provided. The fusion proteins can contain additional components, such as *E. coli* maltose binding protein (MBP) that aid in uptake of the protein by cells (see, International PCT application No. WO 01/32711). Nucleic acid molecules encoding the mutant Rep protein or fusion protein operably linked to a promoter, such as an inducible promoter for expression in mammalian cells are also provided. Such promoters include, but are not

-39-

limited to, CMV and SV40 promoters; adenovirus promoters, such as the E2 gene promoter, which is responsive to the HPV E7 oncoprotein; a PV promoter, such as the PBV p89 promoter that is responsive to the PV E2 protein; and other promoters that are activated by the HIV or PV or 5 oncogenes.

The mutant rep proteins are also delivered to the cells in rAAV or a portion thereof that can additionally encode therapeutic agents for treatment of the cancer or HIV infection or other disorders.

Methods of inhibiting oncogenic transformation by bovine PV (BPV) 10 and by human PV (HPV) are provided.

Methods of inhibiting PV, PV-associated diseases, HIV and HIV-associated diseases are provided. These methods are practiced by administering the proteins, nucleic acids or rAAV or portions thereof to a subject, such as a mammal, including a human to thereby inhibit or 15 modulate disease progression or oncogenic transformation.

#### Other systems

It has been shown that the Rep protein is involved in the regulation of gene expression, including viral replication as described above, cellular pathways and protein phosphorylation (see, e.g., Chiorini *et al.* (1998) 20 *Mol. Cell Biol.* 18:5921-5929). Hence the mutant Rep proteins provided herein can be used to block, stimulate, inhibit, regulate or otherwise modulate metabolic or cellular signaling pathways. Rep proteins provided herein can be used to block, stimulate, inhibit, regulate or otherwise modulate cyclic AMP response pathways, and also to regulate or 25 modulate cellular promoters as a means of modulating gene expression. Methods using these proteins for such purposes are provided herein.

#### Formulation of rAAV

Pharmaceutical compositions containing the rAAV, fusion proteins or encoding nucleic acid molecules can be formulated in any conventional

-40-

manner by mixing a selected amount of rAAV with one or more physiologically acceptable carriers or excipients. For example, the rAAV may be suspended in a carrier such as PBS (phosphate buffered saline). The active compounds can be administered by any appropriate route, for

- 5 example, orally, parenterally, intravenously, intradermally, subcutaneously, or topically, in liquid, semi-liquid or solid form and are formulated in a manner suitable for each route of administration. Preferred modes of administration include oral and parenteral modes of administration.

- 10 The rAAV and physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either through the mouth or the nose) or for oral, buccal, parenteral or rectal administration. For administration by inhalation, the rAAV can be delivered in the form of an aerosol spray presentation from pressurized  
15 packs or a nebulizer, with the use of a suitable propellant, e.g. dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges, e.g. of  
20 gelatin, for use in an inhaler or insufflator may be formulated containing a powder mix of a therapeutic compound and a suitable powder base such as lactose or starch.

- For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional  
25 means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g. magnesium stearate, talc or silica); disintegrants (e.g. potato starch or sodium starch

-41-

glycolate); or wetting agents (e.g. sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for

- 5 constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g. sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g. lecithin or acacia); non-aqueous vehicles (e.g.
- 10 almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g. methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably formulated to

- 15 give controlled release of the active compound. For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

The rAAV may be formulated for parenteral administration by injection e.g. by bolus injection or continuous infusion. Formulations for

- 20 injection may be presented in unit dosage form e.g. in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient
- 25 may be in powder lyophilized form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

In addition to the formulations described previously, the rAAV may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example, subcutaneously or

-42-

intramuscularly) or by intramuscular injection. Thus, for example, the therapeutic compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a

5 sparingly soluble salt.

The active agents may be formulated for local or topical application, such as for topical application to the skin and mucous membranes, such as in the eye, in the form of gels, creams, and lotions and for application to the eye or for intracisternal or intraspinal

10 application. Such solutions, particularly those intended for ophthalmic use, may be formulated as 0.01% - 10% isotonic solutions, pH about 5-7, with appropriate salts. The compounds may be formulated as aerosols for topical application, such as by inhalation (see, e.g., U.S. Patent Nos. 4,044,126, 4,414,209, and 4,364,923, which describe

15 aerosols for delivery of a steroid useful for treatment inflammatory diseases, particularly asthma).

The concentration of active compound in the drug composition will depend on absorption, inactivation and excretion rates of the active compound, the dosage schedule, and amount administered as well as

20 other factors known to those of skill in the art. For example, the amount that is delivered is sufficient to treat the symptoms of hypertension.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example, comprise

25 metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

The active agents may be packaged as articles of manufacture containing packaging material, an agent provided herein, and a label that indicates the disorder for which the agent is provided.

-43-

The following example is included for illustrative purposes only and is not intended to limit the scope of the invention. The specific methods exemplified can be practiced with other species. The examples are intended to exemplify generic processes.

5

## EXAMPLE

### Materials and Methods

#### Cells:

- 293 human embryo kidney (HEK) cells, obtained from ATCC, were cultured in Dulbecco's modified Eagle's medium containing 4.5 g/l glucose (DMEM; GIBCO-BRL) 10 % fetal bovine serum (FBS, Hyclone).  
10 Hela rep-cap 32 cells, described above, were obtained from Anna Salvetti (CHU, Nantes) and cultured in the medium described above.

#### Plasmids:

- pNB-Adeno, which encodes the entire E2A and E4 regions and VA  
15 RNA I and II genes of Adenovirus type 5, was constructed by ligating into the polylinker of multiple cloning site of pBSII KS (+/-) (Stratagene, San Diego, USA) the SalI-HindIII fragment (9842-11555 nt) of Adenovirus type 5 and the BamHI-Clal fragment (21563- 35950) of pBR325. All fragments of adenovirus gene were obtained from the plasmid pBHG-10  
20 (Microbix, Ontario, Canada). pNB-AAV encodes the genes rep and cap of AAV-2 and was constructed by ligation of XbaI-XbaI PCR fragment containing the genome of AAV-2 from nucleotide 200 to 4480 into XbaI site of polylinker MCS of pBSIIKS(+/-). The PCR fragment was obtained from pAV1 (ATCC, USA). Plasmid pNB-AAV was derived from plasmid  
25 pVA1I, which contains the AAV genomic region, rep and cap. pNB-AAV does not contain the AAV ITR's present in pAV1. pAAV-CMV(nls)LacZ was provided by Dr Anna Salvetti (CHU, Nantes).

Plasmid pCMV(nls)LacZ (rAAV vector plasmid) and pNB-Adeno were prepared in DH5 $\alpha$  E.coli and purified by Nucleobond AX PC500 Kit

-44-

(Macherey-Nagel), according to standard procedures. Plasmid pAAV-CMV(nls)LacZ is derived from plasmid psub201 by deleting the rep-cap region with SnaB I and replacing it with an expression cassette harboring the cytomegalovirus (CMV) immediate early promoter (407 bp), the 5 nuclear localized  $\beta$ -galactosidase gene and the bovine growth hormone polyA signal (324 bp) (see, Chadeuf *et al.* (2000) *J. Gene Med.* 2:260-268. pAAV-CMV(nls)LacZ was provided by Dr Anna Salvetti.

**Virus:**

Wild type adenovirus (AV) type 5 stock, originally provided by Dr 10 Philippe Moullier (CHU, Nantes), was produced accordingly to standard procedures.

**Construction of Rep mutant libraries**

25 pmol of each mutagenic primer was placed into a 96 PCR well plate. 15  $\mu$ l of reaction mix (0.25 pmol of pNB-AAV), 25 pmol of the 15 selection primer (changing one non-essential unique restriction site to a new restriction site), 2  $\mu$ l of 10X mutagenesis buffer (100mM Tris-acetate pH7.5, 100 mM MgOAc and 500 mM KOAc pH7.5) was added into each well. The samples were incubated at 98°C for 5 minutes and then immediately incubated for 5 minutes on ice. Finally, the plate was placed 20 at room temperature for 30 minutes.

The primer extension and ligation reactions of the new strands were completed by adding to each sample: 7  $\mu$ l of nucleotide mix (2.86 mM each nucleotide and 1.43 X mutagenesis buffer) and 3 $\mu$ l of a fresh 1:10 enzyme dilution mix (0.025U/ $\mu$ l of native T7 DNA polymerase and 25 1U/ $\mu$ l of T4 DNA ligase were diluted in 20mM Tris HCl pH7.5, 10 mM KCl, 10 mM  $\beta$ - mercaptoethanol, 1 mM DTT, 0.1 mM EDTA and 50% glycerol). Samples were incubated at 37°C for 1 hour. The T4 DNA ligase was inactivated by incubating the reactions at 72°C for 15 minutes to

-45-

prevent re-ligation of the digested strands during the digestion of the parental plasmid (pNB-AAV).

Each mutagenesis reaction was digested with restriction enzyme to eliminate parental plasmids: 30  $\mu$ l solution containing 3 $\mu$ l of 10X enzyme digestion buffer and 10 units of restriction enzyme were added to each mutagenesis reaction and incubated at 37°C for at least 3 hours.

90  $\mu$ l of the *E. coli* XLmutS competent cells (Stratagene, San Diego CA; supplemented with 1.5  $\mu$ l of  $\beta$ -mercaptoethanol to a final concentration of 25 mM) were aliquoted into prechilled deep-well plates.

10 The plates were incubated on ice for 10 minutes and swirling gently every 2 minutes.

A fraction of the reactions that had been digested with restriction enzyme (1/10 of the total volume) was added to the deep well plates. The plates were swirled gently prior to incubation on ice for 30 minutes. A

15 heat pulse was performed in a 42°C water bath for 45 seconds, the transformation mixture was incubated on ice for 2 minutes and 0.45 ml of preheated SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 8.5 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub> and 20 mM glucose at pH 7) was added. The plates were incubated at 37°C for 1 hour with shaking.

20 To enrich for mutant plasmids, 1 ml of 2X YT broth medium (YT medium is 0.5% yeast extract, 0.5% NaCl, 0.8% bacto-tryptone), supplemented with 100  $\mu$ g/ml of ampicillin, was added to each transformation mixture and the cultures were grown overnight at 37°C with shaking. Plasmid DNA isolation was performed from each mutant

25 culture using standard procedure described in Nucleospin Multi-96 Plus Plasmid Kit (Macherey-Nagel). Five hundred  $\mu$ g of the resulting isolated DNA was digested with 10 units of the selection restriction enzyme in a total volume of 30 $\mu$ l containing 3  $\mu$ l of 10X enzyme digestion buffer for overnight at 37°C.

-46-

A fraction of the digested reactions (1/10 of the total volume) were transformed into 40  $\mu$ l of Epicurian coli XL1-Blue competent cells supplemented with 0.68  $\mu$ l of  $\beta$ -mercaptoethanol to a final concentration of 25 mM. After heat pulse, 0.45 ml of SOC was added and the

- 5 transformation mixtures were incubated for 1 hour at 37°C with shaking before to be plate on LB-ampicillin agar plates. The agar plates were incubated overnight at 37°C and the colonies obtained were picked up and grown overnight at 37°C into deep-well plates.

- Four clones per reaction were screened for the presence of the  
10 mutation using restriction enzyme specific to the new restriction site introduced into the mutated plasmid with the selection primer. The cDNA from selected clones was also sequenced to confirm the presence of the expected mutation.

#### Monitoring rAAV Production

- 15 rAAV from each of the above wells, were produced by triple transfection on 293 HEK cells.  $3 \times 10^4$  cells were seeded into each well of 96 micro-well plate and cultured for 24 hours before transfection. Transfection was made on cells at about 70% confluence. 25 kDa PEI (poly-ethylene-imine, Sigma-Aldrich) was used for the triple transfection  
20 step. Equimolar amounts of the three plasmids AV helper plasmid (pNB-Adeno), AAV helper plasmid (pNB-AAV or a mutant clone rep plasmid) and vector plasmid (pAAV-CMV(nls)LacZ) were mixed with 10 mM PEI by gently shaking. The mixture was the added to the medium culture on the cells. 60 hours after transfection, the culture medium was replaced with  
25 100  $\mu$ l of lysis buffer (50mM Hepes, pH 7.4; 150 mM NaCl; 1mM MgCl<sub>2</sub>; 1 mM CaCl<sub>2</sub>; 0.01% CHAPS). After one cycle of freeze-thawing the cellular lysate was filtered through a millipore filter 96 well plate and stored at -80°C.

-47-

**rAAV infection particles (ip)**

Titers of rAAV vector particles were determined on HeLa rep/cap 32 cells using standard dRA (serial dilution replication assay) test. Cells were plated 24 hours before infection at a density of  $1 \times 10^4$  cells in 96-well plates. Serial dilutions of the rAAV preparation were made between 1 and  $1 \times 10^6 \mu\text{l}$  and used for co-infection of the HeLa rep/cap 32 cells together with wt-AV type 5 (MOI 25). 48 hours after infection the ip were measured by real time PCR or by the quantification of biological activity of the transgene.

**10 Real Time PCR**

Infected HeLa rep/cap 32 cells were lysed with  $50 \mu\text{l}$  of solution (50 mM Hepes, pH 7.4; 150 mM NaCl). After one cycle of freeze-thawing  $50 \mu\text{l}$  of Proteinase K (10 mg/ml) and the lysate were incubated one hour at 55°C. The enzyme was inactivated by incubation 10 min at 96°C.

**15** For real time PCR,  $0.2 \mu\text{l}$  of lysate was taken. Final volume of the reaction was  $10 \mu\text{l}$  in 384 well plate using an Applied Biosystem Prism 7900. The primers and fluorescence probe set corresponding to the CMV promoter were as follows: CMV 1 primer 5'-TGCCAAGTACGCCCTAT-3' (SEQ ID No. 733) ( $0.2 \mu\text{M}$ ) and CMV 2 primer 5'-AGGTCATGTACTGGGCATAATGC -3' (SEQ ID No. 734) ( $0.2 \mu\text{M}$ ) ; probe VIC-Tamra 5'-TCAATGACGGTAAATGGCCCGCCT-3' (SEQ ID No. 735) ( $0.1 \mu\text{M}$ ). dRA plots were obtained by plotting the DNA copy number (obtained by real time PCR) vs. the dilution of the rAAV preparation.

**25  $\beta$ -Galactosidase activity**

After 48 hours of infection, cells were treated with trypsin, and  $100 \mu\text{l}$  of reaction solution (GalScreen Kit, Tropix) was added and incubated for one hour at 26 °C. Luminescence was measured in NorthStar (Tropix) HTS station. dRA plots were obtained plotting the

-48-

intensity of  $\beta$ -Galactosidase activity vs. the dilution of the rAAV preparation.

**Mathematical Model for results analysis:**

Results were analyzed using the Hill equation-based analysis

- 5 (designated NautScan™; see, Patent n° 9915884, 1999, France; published as International PCT application No. WO 01/44809 (PCT n° PCT/FR00/03503, Dec, 2000). Briefly, data were processed using a Hill equation-based model that allows extraction of key feature indicators of performance for each individual mutant. Mutants were ranked based on
- 10 the values of their individual performance and those at the top of the ranking list were selected as Leads.

**Results**

**Generation of diversity.**

To identify candidate amino acid (aa) positions on the rep protein

- 15 involved in rep protein activity an Ala-scan was performed on the rep sequence. For this, each amino acid in the rep protein sequence was replaced with Alanine. To do this sets of rAAV that encode mutant rep proteins in which each differs from wild type by replacement of one amino acid with Ala, were generated. Each set of rAAV was individually
- 20 introduced into cells in a well of a microtiter plate, under conditions for expression of the rep protein. The amount of virus that could be produced from each variant was measured as described below. Briefly, activity of Rep was assessed by determining the amount of AAV or rAAV produced using infection assays on HeLa Rep-cap 32 cells and by
- 25 measurement of AAV DNA replication using Real Time PCR, or by assessing transgene ( $\beta$ -galactosidase) expression. The relative activity of each individual mutant compared to the native protein was assessed and "hits" identified. Hit positions are the positions in the mutant proteins that resulted in an alteration (selected to be at least about 20%), in this

-49-

instance all resulted in a decrease, in the amount of virus produced compared to the activity of the native (wildtype) gene (see Fig. 2A).

The hits were then used for identification of leads (see, Fig. 2B). Assays for Rep activity were performed as described for identification of 5 the hit positions. Hit positions on Rep proteins and the effect of specific amino acids on the productivity of AAV-2 summarized in the following table:

	Hit position	replacing amino acid (effect)
10	4 (ttt) F	(gct) A (decrease)
	10 (aag) K	(gcg) A (decrease)
	20 (ccc) P	(gcc) A (decrease)
	22 (att) I	(gct) A (decrease)
	28 (tgg) W	(gcg) A (decrease)
15	32 (gag) E	(gcg) A (decrease)
	38 (ccg) P	(gcg) A (decrease)
	39 (cca) P	(gca) A (decrease)
	54 (ctg) L	(gct) A (decrease)
	59 (ctg) L	(gcg) A (decrease)
20	64 (ctg) L	(gcg) A (decrease)
	74 (ccg) P	(gcg) A (decrease)
	86 (gag) E	(gcg) A (decrease)
	88 (tac) Y	(gcc) A (decrease)
	101 (aaa) K	(gca) A (decrease)
25	124 (atc) I	(gcc) A (decrease)
	125 (gag) E	(gcg) A (decrease)
	127 (act) T	(gct) A (decrease)
	132 (ttc) F	(gcc) A (decrease)
	140 (ggc) G	(gcc) A (decrease)

-50-

	Hit position	replacing amino acid (effect)
	161 (acc) T	(gcc) A (decrease)
	163 (cct) P	(gct) A (decrease)
	175 (tat) Y	(gct) A (decrease)
5	193 (ctg) L	(gcg) A (decrease)
	196 (gtg) V	(gcg) A (decrease)
	197 (tcg) S	(gcc) A (decrease)
	221 (tca) S	(gca) A (decrease)
	228 (gtc) V	(gcg) A (decrease)
	231 (ctc) L	(gcc) A (decrease)
10	234 (aag) K	(gcg) A (decrease)
	237 (acc) T	(gcc) A (decrease)
	250 (tac) Y	(gcc) A (decrease)
	258 (aac) N	(gcc) A (decrease)
	260 (cgg) R	(gcg) A (decrease)
15	263 (atc) I	(gcc) A (decrease)
	264 (aag) K	(gcg) A (decrease)
	334 (ggg) G	(gcg) A (decrease)
	335 (cct) V	(gct) A (decrease)
	337 (act) T	(gct) A (decrease)
20	341 (acc) T	(gcc) A (decrease)
	342 (aac) N	(gcc) A (decrease)
	347 (ata) I	(gca) A (decrease)
	350 (act) T	(gct) A (decrease)
	354 (tac) Y	(gcc) A (decrease)
25	363 (aac) N	(gcc) A (decrease)
	364 (ttt) F	(gct) A (decrease)

-51-

	Hit position	replacing amino acid (effect)	
	367 (aac) N	(gcc) A (decrease)	
	370 (gtc) V	(gcc) A (decrease)	
	376 (tgg) W	(gcg) A (decrease)	
	381 (aag) K	(gcg) A (decrease)	
5	382 (atg) M	(gcg) A (decrease)	
	389 (tcg) S	(gcg) A (decrease)	
	407 (tcc) S	(gcc) A (decrease)	
	411 (ata) I	(gca) A (decrease)	
	414 (act) T	(gct) A (decrease)	
10	420 (tcc) S	(gct) A (decrease)	
	421 (aac) N	(gcc) A (decrease)	
	422 (acc) T	(gcc) A (decrease)	
	424 (atg) M	(gcg) A (decrease)	
	428 (att) I	(gct) A (decrease)	
15	429 (gac) D	(gcc) A (decrease)	
	438 (cag) Q	(gcg) A (decrease)	
	440 (ccg) P	(gcg) A (decrease)	
	451 (acc) T	(gcc) A (decrease)	
	460 (aag) K	(gcg) A (decrease)	
20	462 (acc) T	(gcc) A (decrease)	(ata) I (increase)
	484 (ttc) F	(gcc) A (decrease)	
	488 (aag) K	(gcg) A (decrease)	
	495 (ccc) P	(gcc) A (decrease)	
	497 (ccc) P	(gcc) A (decrease)	(cga) R (increase)
25	497 (ccc) P	(gcc) A (decrease)	(ctc) L (increase)
	497 (ccc) P	(gcc) A (decrease)	(tac) Y (increase)

-52-

	Hit position	replacing amino acid (effect)	
	498 (agt) S	(gct) A (decrease)	
	499 (gac) D	(gcc) A (decrease)	
	503 (agt) S	(gcg) A (decrease)	
	511 (tca) S	(gca) A (decrease)	
5	512 (gtt) V	(gct) A (decrease)	
	516 (tcg) S	(gcg) A (decrease)	
	517 (acg) T	(gct) A (decrease)	(aac) N (increase)
	518 (tca) S	(gca) A (decrease)	
	519 (gac) D	(gcg) A (decrease)	
10	542 (ctg) L	(gcg) A (decrease)	(tcg) S (increase)
	548 (aga) R	(gca) A (decrease)	(agc) S (increase)
	598 (gga) G	(gca) A (decrease)	(agc) S (increase)
	600 (gtg) V	(gcg) A (decrease)	(ccg) P (increase)
	601 (cca) P	(gca) A (decrease)	
15	Hit position (within intron)	replacing sequence (effect)	
	630 (tgc)	gcg (decrease)	cgc or tca or cct (increase)

The hits in other AAV serotypes (see, also Figures 3A and 3B) are  
20 as follows:

	HIT POSITION						
	AAV-2	AAV-1	AAV-3	AAV-3B	AAV-4	AAV-6	AAV-5
	4	4	4	4	4	4	4
	10	10	10	10	10	10	10
25	20	20	20	20	20	20	20
	22	22	22	22	22	22	22
	29	29	29	29	29	29	29

-53-

HIT POSITION							
	32	32	32	32	32	32	32
5	38	38	38	38	38	38	38
	39	39	39	39	39	39	39
	54	54	54	54	54	54	54
	59	59	59	59	59	59	59
	64	64	64	64	64	64	64
	74	74	74	74	74	74	
	86	86	86	86	86	86	85
	88	88	88	88	88	88	87
10	101	101	101	101	101	101	100
	124	124	124	124	124	124	123
	125	125	125	125	125	125	124
	127	127	127	127	127	127	126
	132	132	132	132	132	132	131
15	140	140	140	140	140	140	
	161	161	161	161	161	161	158
	163	163	163	163	163	163	160
	175	175	175	175	175	175	172
	193	193	193	193	193	193	190
20	196	196	196	196	196	196	193
	197	197	197	197	197	197	194
	221	221	221	221	221	221	217
	228	228	228	228	228	228	224
	231	231	231	231	231	231	227
25	234	234	234	234	234	234	230
	237	237	237	237	237	237	233
	250	250	250	250	250	250	246
	258	258	258	258	258	258	254
	260	260	260	260	260	260	256

-54-

HIT POSITION							
	263	263	263	263	263	263	259
5	264	264	264	264	264	264	260
	334	334	334	334	334	334	330
	335	335	335	335	335	335	331
	337	337	337	337	337	337	333
	341	341	341	341	341	341	337
	342	342	342	342	342	342	338
	347	347	347	347	347	347	342
10	350	350	350	350	350	350	346
	354	354	354	354	354	354	350
	363	363	363	363	363	363	359
	364	364	364	364	364	364	360
	367	367	367	367	367	367	363
15	370	370	370	370	370	370	366
	376	376	376	376	376	376	372
	381	381	381	381	381	381	377
	382	382	382	382	382	382	378
20	389	389	389	389	389	389	385
	407	407	407	407	407	407	403
	411	411	411	411	411	411	407
	414	414	414	414	414	414	410
	420	420	420	420	420	420	416
	421	421	421	421	421	421	417
	422	422	422	422	422	422	418
25	424	424	424	424	424	424	420
	428	428	428	428	428	428	424
	429	429	429	429	429	429	425
	438	438	438	438	438	438	434
	440	440	440	440	440	440	436

-55-

HIT POSITION							
451	451	451	451	451	451	451	447
460	460	460	460	460	460	460	456
462	462	462	462	462	462	462	458
484	484	484	484	484	484	484	480
5	488	488	488	488	488	488	484
495	495	495	495	495	495	495	491
497	497	497	497	497	497	497	493
498	498	498	498	498	498	498	494
499	499	499	499	499	499	499	495
10	503	503	503	503	503	503	499
511	511	511	511	511	511	511	529
512	512	512	512	512	512	512	530
516	516	516	516	516	516	516	534
517	517	517	517	517	517	517	535
15	518	518	518	518	518	518	536
519	519	519	519	519	519	519	537
542	543	542	542	542	543	543	561
548	549	548	548	548	549	549	567
598	599	600	600	599	599	599	—
20	600	602	603	603	602	602	589
	601	603	604	604	603	603	590

Sets of nucleic acids encoding the rep protein were generated. The rep proteins encoded by these sets of nucleic acid molecules were those in which each amino acid position identified as a "hit" in the ala-scan step, were each sequentially replaced by all remaining 18 amino acids using site directed mutagenesis. Each mutant was designed, generated, processed and analyzed physically separated from the others in

-56-

addressable arrays. No mixtures, pools, nor combinatorial processing were used.

As in the first round (alanine scan), a library of mutant rAAV was generated in which each individual mutant was independently and

5 individually generated in a independent reaction and such that each mutant contains only a single amino acid change and this for each amino acid residue. Again, each resulting mutant rep protein was then expressed and the amount of virus produced in cells assessed and compared to the native protein.

**10 Lead identification**

Since rep proteins that result in increased virus production are of interest, those mutants that lead to an increase in the amount of virus produced (2 to 10 times the native activity), were selected as "leads." Ten such mutants were identified.

15 Based on the results obtained from the assays described above (i.e. titer of virus produced by each rep variant), each individual rep variant was assigned a specific activity. Those variant proteins displaying the highest titers were selected as leads (see Table above). Leads include: amino acid replacement of T by N at Hit position 350; T by I at Hit  
20 position 462; P by R at Hit position 497; P by L at Hit position 497; P by Y at Hit position 497; T by N at Hit position 517; L by S at Hit position 542; R by S at Hit positio 548, G by S at Hit position 598; G by D at Hit position 598; V by P at Hit position 600.

Also provided are combinations of the above mutant Rep 78, 68,  
25 52. 40 proteins, nucleic acids encoding the proteins, and recombinant AAV (any serotype) containing the mutation at the indicated position or corresponding position for serotypes other than AAV-2, including any set forth in the following table and corresponding SEQ ID Nos. Each amino acid sequence is set forth in a separate sequence ID listing; for each

-57-

mutation or combination thereof there is a single SEQ ID setting forth the unspliced nucleic acid sequence for Rep78/68, which for all mutations from amino acid 228 on, includes the corresponding Rep 52 and Rep 40 encoding sequence as well.

#### 5 Amino acid sequences of exemplary mutant Rep proteins

	Seq no.	gene	position(s)	codon(s)
10	seq.1	rep78	4	GCT
	seq.2	rep68	4	GCT
	seq.3	rep78	10	GCG
	seq.4	rep68	10	GCG
	seq.5	rep78	20	GCC
	seq.6	rep68	20	GCC
	seq.7	rep78	22	GCT
	seq.8	rep68	22	GCT
15	seq.9	rep78	29	GCG
	seq.10	rep68	29	GCG
	seq.11	rep78	38	GCG
	seq.12	rep68	38	GCG
	seq.13	rep78	39	GCA
20	seq.14	rep68	39	GCA
	seq.15	rep78	53	GCT
	seq.16	rep68	53	GCT
	seq.17	rep78	59	GCG
25	seq.18	rep68	59	GCG
	seq.19	rep78	64	GCT
	seq.20	rep68	64	GCT
	seq.21	rep78	74	GCG
30	seq.22	rep68	74	GCG
	seq.23	rep78	86	GCG
	seq.24	rep68	86	GCG
	seq.25	rep78	88	GCC
	seq.26	rep68	88	GCC
35	seq.27	rep78	101	GCA
	seq.28	rep68	101	GCA
	seq.29	rep78	124	GCC
	seq.30	rep68	124	GCC
40	seq.31	rep78	125	GCG
	seq.32	rep68	125	GCG
	seq.33	rep78	127	GCT
	seq.34	rep68	127	GCT
45	seq.35	rep78	132	GCC
	seq.36	rep68	132	GCC
	seq.37	rep78	140	GCC
	seq.38	rep68	140	GCC
45	seq.39	rep78	161	GCC
	seq.40	rep68	161	GCC

-58-

	seq.41	rep78	163	GCT
	seq.42	rep68	163	GCT
	seq.43	rep78	175	GCT
	seq.44	rep68	175	GCT
5	seq.45	rep78	193	GCG
	seq.46	rep68	193	GCG
	seq.47	rep78	196	GCC
	seq.48	rep68	196	GCC
	seq.49	rep78	197	GCC
10	seq.50	rep68	197	GCC
	seq.51	rep78	221	GCA
	seq.52	rep68	221	GCA
	seq.53	rep78	228	GCG
	seq.54	rep52	228	GCG
15	seq.55	rep68	228	GCG
	seq.56	rep40	228	GCG
	seq.57	rep78	231	GCC
	seq.58	rep52	231	GCC
	seq.59	rep68	231	GCC
20	seq.60	rep40	231	GCC
	seq.61	rep78	234	GCG
	seq.62	rep52	234	GCG
	seq.63	rep68	234	GCG
	seq.64	rep40	234	GCG
25	seq.65	rep78	237	GCC
	seq.66	rep52	237	GCC
	seq.67	rep68	237	GCC
	seq.68	rep40	237	GCC
	seq.69	rep78	250	GCC
30	seq.70	rep52	250	GCC
	seq.71	rep68	250	GCC
	seq.72	rep40	250	GCC
	seq.73	rep78	258	GCC
	seq.74	rep52	258	GCC
35	seq.75	rep68	258	GCC
	seq.76	rep40	258	GCC
	seq.77	rep78	260	GCG
	seq.78	rep52	260	GCG
	seq.79	rep68	260	GCG
40	seq.80	rep40	260	GCG
	seq.81	rep78	263	GCC
	seq.82	rep52	263	GCC
	seq.83	rep68	263	GCC
	seq.84	rep40	263	GCC
45	seq.85	rep78	264	GCG
	seq.86	rep52	264	GCG
	seq.87	rep68	264	GCG
	seq.88	rep40	264	GCG
	seq.89	rep78	334	GCG
50	seq.90	rep52	334	GCG
	seq.91	rep68	334	GCG

-59-

	seq.92	rep40	334	GCG
	seq.93	rep78	335	GCT
	seq.94	rep52	335	GCT
	seq.95	rep68	335	GCT
5	seq.96	rep40	335	GCT
	seq.97	rep78	337	GCT
	seq.98	rep52	337	GCT
	seq.99	rep68	337	GCT
	seq.100	rep40	337	GCT
10	seq.101	rep78	341	GCC
	seq.102	rep52	341	GCC
	seq.103	rep68	341	GCC
	seq.104	rep40	341	GCC
	seq.105	rep78	342	GCC
15	seq.106	rep52	342	GCC
	seq.107	rep68	342	GCC
	seq.108	rep40	342	GCC
	seq.109	rep78	347	GCA
	seq.110	rep52	347	GCA
20	seq.111	rep68	347	GCA
	seq.112	rep40	347	GCA
	seq.113	rep78	350	AAT
	seq.114	rep52	350	AAT
	seq.115	rep68	350	AAT
25	seq.116	rep40	350	AAT
	seq.117	rep78	350	GCT
	seq.118	rep52	350	GCT
	seq.119	rep68	350	GCT
	seq.120	rep40	350	GCT
30	seq.121	rep78	354	GCC
	seq.122	rep52	354	GCC
	seq.123	rep68	354	GCC
	seq.124	rep40	354	GCC
	seq.125	rep78	363	GCC
35	seq.126	rep52	363	GCC
	seq.127	rep68	363	GCC
	seq.128	rep40	363	GCC
	seq.129	rep78	364	GCT
	seq.130	rep52	364	GCT
40	seq.131	rep68	364	GCT
	seq.132	rep40	364	GCT
	seq.133	rep78	367	GCC
	seq.134	rep52	367	GCC
	seq.135	rep68	367	GCC
45	seq.136	rep40	367	GCC
	seq.137	rep78	370	GCC
	seq.138	rep52	370	GCC
	seq.139	rep68	370	GCC
	seq.140	rep40	370	GCC
50	seq.141	rep78	376	GCG
	seq.142	rep52	376	GCG

-60-

	seq.143	rep68	376	GCG
	seq.144	rep40	376	GCG
	seq.145	rep78	381	GCG
	seq.146	rep52	381	GCG
5	seq.147	rep68	381	GCG
	seq.148	rep40	381	GCG
	seq.149	rep78	382	GCG
	seq.150	rep52	382	GCG
	seq.151	rep68	382	GCG
10	seq.152	rep40	382	GCG
	seq.153	rep78	389	GCG
	seq.154	rep52	389	GCG
	seq.155	rep68	389	GCG
	seq.156	rep40	389	GCG
15	seq.157	rep78	407	GCC
	seq.158	rep52	407	GCC
	seq.159	rep68	407	GCC
	seq.160	rep40	407	GCC
	seq.161	rep78	411	GCA
20	seq.162	rep52	411	GCA
	seq.163	rep68	411	GCA
	seq.164	rep40	411	GCA
	seq.165	rep78	414	GCT
	seq.166	rep52	414	GCT
25	seq.167	rep68	414	GCT
	seq.168	rep40	414	GCT
	seq.169	rep78	420	GCT
	seq.170	rep52	420	GCT
	seq.171	rep68	420	GCT
30	seq.172	rep40	420	GCT
	seq.173	rep78	421	GCC
	seq.174	rep52	421	GCC
	seq.175	rep68	421	GCC
	seq.176	rep40	421	GCC
35	seq.177	rep78	422	GCC
	seq.178	rep52	422	GCC
	seq.179	rep68	422	GCC
	seq.180	rep40	422	GCC
	seq.181	rep78	424	GCG
40	seq.182	rep52	424	GCG
	seq.183	rep68	424	GCG
	seq.184	rep40	424	GCG
	seq.185	rep78	428	GCT
	seq.186	rep52	428	GCT
45	seq.187	rep68	428	GCT
	seq.188	rep40	428	GCT
	seq.189	rep78	429	GCC
	seq.190	rep52	429	GCC
	seq.191	rep68	429	GCC
50	seq.192	rep40	429	GCC
	seq.193	rep78	438	GCG

-61-

	seq.194	rep52	438	GCG
	seq.195	rep68	438	GCG
	seq.196	rep40	438	GCG
	seq.197	rep78	440	GCG
5	seq.198	rep52	440	GCG
	seq.199	rep68	440	GCG
	seq.200	rep40	440	GCG
	seq.201	rep78	451	GCC
	seq.202	rep52	451	GCC
10	seq.203	rep68	451	GCC
	seq.204	rep40	451	GCC
	seq.205	rep78	460	GCG
	seq.206	rep52	460	GCG
	seq.207	rep68	460	GCG
15	seq.208	rep40	460	GCG
	seq.209	rep78	462	GCC
	seq.210	rep52	462	GCC
	seq.211	rep68	462	GCC
	seq.212	rep40	462	GCC
20	seq.213	rep78	462	ATA
	seq.214	rep52	462	ATA
	seq.215	rep68	462	ATA
	seq.216	rep40	462	ATA
	seq.217	rep78	484	GCC
25	seq.218	rep52	484	GCC
	seq.219	rep68	484	GCC
	seq.220	rep40	484	GCC
	seq.221	rep78	488	GCG
	seq.222	rep52	488	GCG
30	seq.223	rep68	488	GCG
	seq.224	rep40	488	GCG
	seq.225	rep78	495	GCC
	seq.226	rep52	495	GCC
	seq.227	rep68	495	GCC
35	seq.228	rep40	495	GCC
	seq.229	rep78	497	GCC
	seq.230	rep52	497	GCC
	seq.231	rep68	497	GCC
	seq.232	rep40	497	GCC
40	seq.233	rep78	497	CGA
	seq.234	rep52	497	CGA
	seq.235	rep68	497	CGA
	seq.236	rep40	497	CGA
	seq.237	rep78	497	CTC
45	seq.238	rep52	497	CTC
	seq.239	rep68	497	CTC
	seq.240	rep40	497	CTC
	seq.241	rep78	497	TAC
	seq.242	rep52	497	TAC
50	seq.243	rep68	497	TAC
	seq.244	rep40	497	TAC

-62-

	seq.245	rep78	498	GCT
	seq.246	rep52	498	GCT
	seq.247	rep68	498	GCT
	seq.248	rep40	498	GCT
5	seq.249	rep78	499	GCC
	seq.250	rep52	499	GCC
	seq.251	rep68	499	GCC
	seq.252	rep40	499	GCC
	seq.253	rep78	503	GCG
10	seq.254	rep52	503	GCG
	seq.255	rep68	503	GCG
	seq.256	rep40	503	GCG
	seq.257	rep78	510	GCA
	seq.258	rep52	510	GCA
15	seq.259	rep68	510	GCA
	seq.260	rep40	510	GCA
	seq.261	rep78	511	GCA
	seq.262	rep52	511	GCA
	seq.263	rep68	511	GCA
20	seq.264	rep40	511	GCA
	seq.265	rep78	512	GCT
	seq.266	rep52	512	GCT
	seq.267	rep68	512	GCT
	seq.268	rep40	512	GCT
25	seq.269	rep78	516	GCG
	seq.270	rep52	516	GCG
	seq.271	rep68	516	GCG
	seq.272	rep40	516	GCG
	seq.273	rep78	517	GCT
30	seq.274	rep52	517	GCT
	seq.275	rep68	517	GCT
	seq.276	rep40	517	GCT
	seq.277	rep78	517	AAC
	seq.278	rep52	517	AAC
35	seq.279	rep68	517	AAC
	seq.280	rep40	517	AAC
	seq.281	rep78	518	GCA
	seq.282	rep52	518	GCA
	seq.283	rep68	518	GCA
40	seq.284	rep40	518	GCA
	seq.285	rep78	519	GCG
	seq.286	rep52	519	GCG
	seq.287	rep68	519	GCG
	seq.288	rep40	519	GCG
45	seq.289	rep78	598	GCA
	seq.290	rep52	598	GCA
	seq.291	rep78	598	GAC
	seq.292	rep52	598	GAC
	seq.293	rep78	598	AGC
50	seq.294	rep52	598	AGC
	seq.295	rep78	600	GCG

-63-

	seq.296	rep52	600	GCG
	seq.297	rep78	600	CCG
	seq.298	rep52	600	CCG
	seq.299	rep78	601	GCA
5	seq.300	rep52	601	GCA
	seq.301	rep78	335 420 495	GCT GCC GCC
	seq.302	rep52	335 420 495	GCT GCC GCC
	seq.303	rep68	335 420 495	GCT GCC GCC
	seq.304	rep40	335 420 495	GCT GCC GCC
10	seq.305	rep78	39 140	GCA GCC
	seq.306	rep68	39 140	GCA GCC
	seq.307	rep78	279 428 451	GCC GCT GCC
	seq.308	rep52	279 428 451	GCC GCT GCC
	seq.309	rep68	279 428 451	GCC GCT GCC
15	seq.310	rep40	279 428 451	GCC GCT GCC
	seq.311	rep78	125 237 600	GCG GCC GCG
	seq.312	rep52	125 237 600	GCG GCC GCG
	seq.313	rep68	125 237 600	GCG GCC GCG
	seq.314	rep40	125 237 600	GCG GCC GCG
20	seq.315	rep78	163 259	GCT GCG
	seq.316	rep52	163 259	GCT GCG
	seq.317	rep68	163 259	GCT GCG
	seq.318	rep40	163 259	GCT GCG
	seq.319	rep78	17 127 189	GCG GCT GCG
25	seq.320	rep68	17 127 189	GCG GCT GCG
	seq.321	rep78	350 428	GCT GCT
	seq.322	rep52	350 428	GCT GCT
	seq.323	rep68	350 428	GCT GCT
	seq.324	rep40	350 428	GCT GCT
30	seq.325	rep78	54 338 495	GCC GCC GCC
	seq.326	rep52	54 338 495	GCC GCC GCC
	seq.327	rep68	54 338 495	GCC GCC GCC
	seq.328	rep40	54 338 495	GCC GCC GCC
	seq.329	rep78	350 420	GCT GCC
35	seq.330	rep52	350 420	GCT GCC
	seq.331	rep68	350 420	GCT GCC
	seq.332	rep40	350 420	GCT GCC
	seq.333	rep78	189 197 518	GCG GCG GCA
	seq.334	rep52	189 197 518	GCG GCG GCA
40	seq.335	rep68	189 197 518	GCG GCG GCA
	seq.336	rep40	189 197 518	GCG GCG GCA
	seq.337	rep78	468 516	GCC GCG
	seq.338	rep52	468 516	GCC GCG
	seq.339	rep68	468 516	GCC GCG
45	seq.340	rep40	468 516	GCC GCG
	seq.341	rep78	127 221 350 54 140	GCT GCA GCT GCC GCC
	seq.342	rep52	127 221 350 54 140	GCT GCA GCT GCC GCC
	seq.343	rep68	127 221 350 54 140	GCT GCA GCT GCC GCC
	seq.344	rep40	127 221 350 54 140	GCT GCA GCT GCC GCC
50	seq.345	rep78	221 285	GCA GCG
	seq.346	rep52	221 285	GCA GCG

-64-

	seq.347	rep68	221 285	GCA GCG
	seq.348	rep40	221 285	GCA GCG
	seq.349	rep78	23 495	GCT GCC
	seq.350	rep52	23 495	GCT GCC
5	seq.351	rep68	23 495	GCT GCC
	seq.352	rep40	23 495	GCT GCC
	seq.353	rep78	20 54 420 495	GCC GCC GCC GCC
	seq.354	rep52	20 54 420 495	GCC GCC GCC GCC
	seq.355	rep68	20 54 420 495	GCC GCC GCC GCC
10	seq.356	rep40	20 54 420 495	GCC GCC GCC GCC
	seq.357	rep78	412 612	GCC GCG
	seq.358	rep52	412 612	GCC GCG
	seq.359	rep68	412 612	GCC GCG
	seq.360	rep40	412 612	GCC GCG
15	seq.361	rep78	197 412	GCG GCC
	seq.362	rep52	197 412	GCG GCC
	seq.363	rep68	197 412	GCG GCC
	seq.364	rep40	197 412	GCG GCC
	seq.365	rep78	412 495 511	GCC GCC GCA
20	seq.366	rep52	412 495 511	GCC GCC GCA
	seq.367	rep68	412 495 511	GCC GCC GCA
	seq.368	rep40	412 495 511	GCC GCC GCA
	seq.369	rep78	98 422	GCC GCC
	seq.370	rep52	98 422	GCC GCC
25	seq.371	rep68	98 422	GCC GCC
	seq.372	rep40	98 422	GCC GCC
	seq.373	rep78	17 127 189	GCG GCT GCG
	seq.374	rep68	17 127 189	GCG GCT GCG
	seq.375	rep78	20 54 495	GCC GCC GCC
30	seq.376	rep52	20 54 495	GCC GCC GCC
	seq.377	rep68	20 54 495	GCC GCC GCC
	seq.378	rep40	20 54 495	GCC GCC GCC
	seq.379	rep78	259 54	GCG GCC
	seq.380	rep52	259 54	GCG GCC
35	seq.381	rep68	259 54	GCG GCC
	seq.382	rep40	259 54	GCG GCC
	seq.383	rep78	335 399	GCT GCG
	seq.384	rep52	335 399	GCT GCG
	seq.385	rep68	335 399	GCT GCG
40	seq.386	rep40	335 399	GCT GCG
	seq.387	rep78	221 432	GCA GCA
	seq.388	rep52	221 432	GCA GCA
	seq.389	rep68	221 432	GCA GCA
	seq.390	rep40	221 432	GCA GCA
45	seq.391	rep78	259 516	GCG GCG
	seq.392	rep52	259 516	GCG GCG
	seq.393	rep68	259 516	GCG GCG
	seq.394	rep40	259 516	GCG GCG
	seq.395	rep78	495 516	GCC GCG
50	seq.396	rep52	495 516	GCC GCG
	seq.397	rep68	495 516	GCC GCG

-65-

	seq.398	rep40	495 516	GCC GCG
	seq.399	rep78	414 14	GCT GCC
	seq.400	rep52	414 14	GCT GCC
	seq.401	rep68	414 14	GCT GCC
5	seq.402	rep40	414 14	GCT GCC
	seq.403	rep78	74 402 495	GCG GCC GCC
	seq.404	rep52	74 402 495	GCG GCC GCC
	seq.405	rep68	74 402 495	GCG GCC GCC
	seq.406	rep40	74 402 495	GCG GCC GCC
10	seq.407	rep78	228 462 497	GCC GCC GCC
	seq.408	rep52	228 462 497	GCC GCC GCC
	seq.409	rep68	228 462 497	GCC GCC GCC
	seq.410	rep40	228 462 497	GCC GCC GCC
	seq.411	rep78	290 338	GCG GCC
15	seq.412	rep52	290 338	GCG GCC
	seq.413	rep68	290 338	GCG GCC
	seq.414	rep40	290 338	GCG GCC
	seq.415	rep78	140 511	GCC GCA
	seq.416	rep52	140 511	GCC GCA
20	seq.417	rep68	140 511	GCC GCA
	seq.418	rep40	140 511	GCC GCA
	seq.419	rep78	86 378	GCG GCG
	seq.420	rep52	86 378	GCG GCG
	seq.421	rep68	86 378	GCG GCG
25	seq.422	rep40	86 378	GCG GCG
	seq.423	rep78	54 86	GCC GCG
	seq.424	rep68	54 86	GCC GCG
	seq.425	rep78	54 86	GCC GCG
	seq.426	rep68	54 86	GCC GCG
30	seq.427	rep78	214 495 140	GCG GCC GCC
	seq.428	rep52	214 495 140	GCG GCC GCC
	seq.429	rep68	214 495 140	GCG GCC GCC
	seq.430	rep40	214 495 140	GCG GCC GCC
	seq.431	rep78	495 511	GCC GCA
35	seq.432	rep52	495 511	GCC GCA
	seq.433	rep68	495 511	GCC GCA
	seq.434	rep40	495 511	GCC GCA
	seq.435	rep78	495 54	GCC GCC
	seq.436	rep52	495 54	GCC GCC
40	seq.437	rep68	495 54	GCC GCC
	seq.438	rep40	495 54	GCC GCC
	seq.439	rep78	197 495	GCG GCC
	seq.440	rep52	197 495	GCG GCC
	seq.441	rep68	197 495	GCG GCC
45	seq.442	rep40	197 495	GCG GCC
	seq.443	rep78	261 20	GCC GCC
	seq.444	rep52	261 20	GCC GCC
	seq.445	rep68	261 20	GCC GCC
	seq.446	rep40	261 20	GCC GCC
50	seq.447	rep78	54 20	GCC GCC
	seq.448	rep68	54 20	GCC GCC

-66-

	seq.449	rep78	197 420	GCG GCC
	seq.450	rep52	197 420	GCG GCC
	seq.451	rep68	197 420	GCG GCC
	seq.452	rep40	197 420	GCG GCC
5	seq.453	rep78	54 338 495	GCC GCC GCC
	seq.454	rep52	54 338 495	GCC GCC GCC
	seq.455	rep68	54 338 495	GCC GCC GCC
	seq.456	rep40	54 338 495	GCC GCC GCC
	seq.457	rep78	197 427	GCG GCG
	10 seq.458	rep52	197 427	GCG GCG
	seq.459	rep68	197 427	GCG GCG
15	seq.460	rep40	197 427	GCG GCG
	seq.461	rep78	54 228 370 387	GCC GCC GCC GCG
	seq.462	rep52	54 228 370 387	GCC GCC GCC GCG
	15 seq.463	rep68	54 228 370 387	GCC GCC GCC GCG
	seq.464	rep40	54 228 370 387	GCC GCC GCC GCG
	seq.465	rep78	221 289	GCA GCC
	seq.466	rep52	221 289	GCA GCC
20	seq.467	rep68	221 289	GCA GCC
	seq.468	rep40	221 289	GCA GCC
	seq.469	rep78	54 163	GCC GCT
	seq.470	rep68	54 163	GCC GCT
	seq.471	rep78	341 407 420	GCC GCC GCC
	seq.472	rep52	341 407 420	GCC GCC GCC
	25 seq.473	rep68	341 407 420	GCC GCC GCC
30	seq.474	rep40	341 407 420	GCC GCC GCC
	seq.475	rep78	54 228	GCC GCC
	seq.476	rep52	54 228	GCC GCC
	seq.477	rep68	54 228	GCC GCC
	30 seq.478	rep40	54 228	GCC GCC
	seq.479	rep78	96 125 511	GCA GCG GCA
	seq.480	rep52	96 125 511	GCA GCG GCA
35	seq.481	rep68	96 125 511	GCA GCG GCA
	seq.482	rep40	96 125 511	GCA GCG GCA
	seq.483	rep78	54 163	GCC GCT
	seq.484	rep68	54 163	GCC GCT
	seq.485	rep78	197 420	GCG GCC
	seq.486	rep52	197 420	GCG GCC
	seq.487	rep68	197 420	GCG GCC
40	40 seq.488	rep40	197 420	GCG GCC
	seq.489	rep78	334 428 499	GCG GCT GCC
	seq.490	rep52	334 428 499	GCG GCT GCC
	seq.491	rep68	334 428 499	GCG GCT GCC
	seq.492	rep40	334 428 499	GCG GCT GCC
	45 seq.493	rep78	197 414	GCG GCT
	seq.494	rep52	197 414	GCG GCT
50	seq.495	rep68	197 414	GCG GCT
	seq.496	rep40	197 414	GCG GCT
	seq.497	rep78	30 54 127	GCG GCC GCT
	seq.498	rep68	30 54 127	GCG GCC GCT
	seq.499	rep78	29 260	GCG GCG

-67-

	seq.500	rep52	29 260	GCG GCG
	seq.501	rep68	29 260	GCG GCG
	seq.502	rep40	29 260	GCG GCG
	seq.503	rep78	4 484	GCT GCC
5	seq.504	rep52	4 484	GCT GCC
	seq.505	rep68	4 484	GCT GCC
	seq.506	rep40	4 484	GCT GCC
	seq.507	rep78	258 124 132	GCC GCC GCC
	seq.508	rep52	258 124 132	GCC GCC GCC
10	seq.509	rep68	258 124 132	GCC GCC GCC
	seq.510	rep40	258 124 132	GCC GCC GCC
	seq.511	rep78	231 497	GCC GCC
	seq.512	rep52	231 497	GCC GCC
	seq.513	rep68	231 497	GCC GCC
15	seq.514	rep40	231 497	GCC GCC
	seq.515	rep78	221 258	GCA GCC
	seq.516	rep52	221 258	GCA GCC
	seq.517	rep68	221 258	GCA GCC
	seq.518	rep40	221 258	GCA GCC
20	seq.519	rep78	234 264 326	GCG GCG GCC
	seq.520	rep52	234 264 326	GCG GCG GCC
	seq.521	rep68	234 264 326	GCG GCG GCC
	seq.522	rep40	234 264 326	GCG GCG GCC
	seq.523	rep78	153 398	AGC GCG
25	seq.524	rep52	153 398	AGC GCG
	seq.525	rep68	153 398	AGC GCG
	seq.526	rep40	153 398	AGC GCG
	seq.527	rep78	53 216	GCG GCC
	seq.528	rep68	53 216	GCG GCC
30	seq.529	rep78	22 382	GCT GCG
	seq.530	rep52	22 382	GCT GCG
	seq.531	rep68	22 382	GCT GCG
	seq.532	rep40	22 382	GCT GCG
	seq.533	rep78	231 411	GCC GCA
35	seq.534	rep52	231 411	GCC GCA
	seq.535	rep68	231 411	GCC GCA
	seq.536	rep40	231 411	GCC GCA
	seq.537	rep78	59 305	GCG GCC
	seq.538	rep52	59 305	GCG GCC
40	seq.539	rep68	59 305	GCG GCC
	seq.540	rep40	59 305	GCG GCC
	seq.541	rep78	53 231	GCG GCC
	seq.542	rep52	53 231	GCG GCC
	seq.543	rep68	53 231	GCG GCC
45	seq.544	rep40	53 231	GCG GCC
	seq.545	rep78	258 498	GCC GCT
	seq.546	rep52	258 498	GCC GCT
	seq.547	rep68	258 498	GCC GCT
	seq.548	rep40	258 498	GCC GCT
50	seq.549	rep78	88 231	GCC GCC
	seq.550	rep52	88 231	GCC GCC

-68-

	seq.551	rep68	88 231	GCC GCC
	seq.552	rep40	88 231	GCC GCC
	seq.553	rep78	101 363	GCA GCC
	seq.554	rep52	101 363	GCA GCC
5	seq.555	rep68	101 363	GCA GCC
	seq.556	rep40	101 363	GCA GCC
	seq.557	rep78	354 132	GCC GCC
	seq.558	rep52	354 132	GCC GCC
	seq.559	rep68	354 132	GCC GCC
10	seq.560	rep40	354 132	GCC GCC
	seq.561	rep78	10 132	GCG GCC
	seq.562	rep68	10 132	GCG GCC

**DNA Sequences**

	Sequence	aa position	codon
15	seq.563	4	GCT
	seq.564	10	GCG
	seq.565	20	GCC
	seq.566	22	GCT
	seq.567	29	GCG
20	seq.568	38	GCG
	seq.569	39	GCA
	seq.570	53	GCT
	seq.571	59	GCG
	seq.572	64	GCT
25	seq.573	74	GCG
	seq.574	86	GCG
	seq.575	88	GCC
	seq.576	101	GCA
	seq.577	124	GCC
30	seq.578	125	GCG
	seq.579	127	GCT
	seq.580	132	GCC
	seq.581	140	GCC
	seq.582	161	GCC
35	seq.583	163	GCT
	seq.584	175	GCT
	seq.585	193	GCG
	seq.586	196	GCC
	seq.587	197	GCC
40	seq.588	221	GCA
	seq.589	228 (Rep78/68)	GCG
		228 (Rep52)	GCG
		228 (Rep 40)	GCG
	seq.590	231 (Rep78/68)	GCC
45		231 (Rep 52)	GCC
		231 (Rep 40)	GCC
	seq.591	234 (Rep78/68)	GCG
		234 (Rep 52)	GCG
		234 (Rep 40)	GCG
50	seq.592	237 (Rep78/68)	GCC

-69-

		237 (Rep 52)	GCC
		237 (Rep 40)	GCC
	seq.593	250 (Rep78/68)	GCC
		250	GCC
5		250	GCC
	seq.594	258 (Rep78/68)	GCC
		258	GCC
		258	GCC
	seq.595	260 (Rep78/68)	GCG
10		260	GCG
		260	GCG
	seq.596	263 (Rep78/68)	GCC
		263	GCC
		263	GCC
15	seq.597	264 (Rep78/68)	GCG
		264	GCG
		264	GCG
	seq.598	334 (Rep78/68)	GCG
		334	GCG
20		334	GCG
	seq.599	335 (Rep78/68)	GCT
		335	GCT
		335	GCT
	seq.600	337 (Rep78/68)	GCT
25		337	GCT
		337	GCT
	seq.601	341 (Rep78/68)	GCC
		341	GCC
		341	GCC
30	seq.602	342 (Rep78/68)	GCC
		342	GCC
		342	GCC
	seq.603	347 (Rep78/68)	GCA
		347	GCA
35		347	GCA
	seq.604	350 (Rep78/68)	AAT
		350	AAT
		350	AAT
	seq.605	350 (Rep78/68)	GCT
40		350	GCT
		350	GCT
	seq.606	354 (Rep78/68)	GCC
		354	GCC
		354	GCC
45	seq.607	363 (Rep78/68)	GCC
		363	GCC
		363	GCC
	seq.608	364 (Rep78/68)	GCT
		364	GCT
50		364	GCT
	seq.609	367 (Rep78/68)	GCC

-70-

		367	GCC
		367	GCC
	seq.610	370 (Rep78/68)	GCC
		370	GCC
5		370	GCC
	seq.611	376 (Rep78/68)	GCG
		376	GCG
		376	GCG
	seq.612	381 (Rep78/68)	GCG
10		381	GCG
		381	GCG
	seq.613	382 (Rep78/68)	GCG
		382	GCG
		382	GCG
15	seq.614	389 (Rep78/68)	GCG
		389	GCG
		389	GCG
	seq.615	407 (Rep78/68)	GCC
		407	GCC
20		407	GCC
	seq.616	411 (Rep78/68)	GCA
		411	GCA
		411	GCA
25	seq.617	414 (Rep78/68)	GCT
		414	GCT
		414	GCT
	seq.618	420 (Rep78/68)	GCT
		420	GCT
		420	GCT
30	seq.619	421 (Rep78/68)	GCC
		421	GCC
		421	GCC
	seq.620	422 (Rep78/68)	GCC
		422	GCC
35		422	GCC
	seq.621	424 (Rep78/68)	GCG
		424	GCG
		424	GCG
40	seq.622	428 (Rep78/68)	GCT
		428	GCT
		428	GCT
	seq.623	429 (Rep78/68)	GCC
		429	GCC
		429	GCC
45	seq.624	438 (Rep78/68)	GCG
		438	GCG
		438	GCG
	seq.625	440 (Rep78/68)	GCG
		440	GCG
50		440	GCG
	seq.626	451 (Rep78/68)	GCC

-71-

		451	GCC
		451	GCC
	seq.627	460 (Rep78/68)	GCG
		460	GCG
5		460	GCG
	seq.628	462 (Rep78/68)	GCC
		462	GCC
		462	GCC
10	seq.629	462 (Rep78/68)	ATA
		462	ATA
		462	ATA
	seq.630	484 (Rep78/68)	GCC
		484	GCC
		484	GCC
15	seq.631	488 (Rep78/68)	GCG
		488	GCG
		488	GCG
	seq.632	495 (Rep78/68)	GCC
		495	GCC
20		495	GCC
	seq.633	497 (Rep78/68)	GCC
		497	GCC
		497	GCC
25	seq.634	497 (Rep78/68)	CGA
		497	CGA
		497	CGA
	seq.635	497 (Rep78/68)	CTC
		497	CTC
		497	CTC
30	seq.636	497 (Rep78/68)	TAC
		497	TAC
		497	TAC
	seq.637	498 (Rep78/68)	GCT
		498	GCT
35		498	GCT
	seq.638	499 (Rep78/68)	GCC
		499	GCC
		499	GCC
40	seq.639	503 (Rep78/68)	GCG
		503	GCG
		503	GCG
	seq.640	510 (Rep78/68)	GCA
		510	GCA
		510	GCA
45	seq.641	511 (Rep78/68)	GCA
		511	GCA
		511	GCA
	seq.642	512 (Rep78/68)	GCT
		512	GCT
50		512	GCT
	seq.643	516 (Rep78/68)	GCG

-72-

		516	GCG
		516	GCG
	seq.644	517 (Rep78/68)	GCT
		517	GCT
5	seq.645	517 (Rep78/68)	AAC
		517	AAC
	seq.646	518 (Rep78/68)	GCA
10		518	GCA
		518	GCA
	seq.647	519 (Rep78/68)	GCG
		519	GCG
		519	GCG
15	seq.648	598 (Rep78/68)	GCA
	seq.649	600 (Rep78/68)	GCG
	seq.650	601 (Rep78/68)	GCA
	seq.651	335 420 495	GCT GCC GCC
		335 420 495	GCT GCC GCC
20		335 420 495	GCT GCC GCC
	seq.652	39 140	GCA GCC
	seq.653	279 428 451	GCC GCT GCC
		279 428 451	GCC GCT GCC
		279 428 451	GCC GCT GCC
25	seq.654	125 237 600	GCG GCC GCG
		125 237 600	GCG GCC GCG
		125 237 600	GCG GCC GCG
	seq.655	163 259	GCT GCG
		163 259	GCT GCG
30		163 259	GCT GCG
	seq.656	17 127 189	GCG GCT GCG
	seq.657	350 428	GCT GCT
		350 428	GCT GCT
		350 428	GCT GCT
35	seq.658	54 338 495	GCC GCC GCC
		54 338 495	GCC GCC GCC
		54 338 495	GCC GCC GCC
	seq.659	350 420	GCT GCC
		350 420	GCT GCC
		350 420	GCT GCC
40		350 420	GCT GCC
	seq.660	189 197 518	GCG GCG GCA
		189 197 518	GCG GCG GCA
		189 197 518	GCG GCG GCA
	seq.661	468 516	GCC GCG
45		468 516	GCC GCG
		468 516	GCC GCG
	seq.662	127 221 350 54 140	GCT GCA GCT GCC GCC
		127 221 350 54 140	GCT GCA GCT GCC GCC
		127 221 350 54 140	GCT GCA GCT GCC GCC
50	seq.663	221 285	GCA GCG
		221 285	GCA GCG

-73-

		221 285	GCA GCG
	seq.664	23 495	GCT GCC
		23 495	GCT GCC
		23 495	GCT GCC
5	seq.665	20 54 420 495	GCC GCC GCC GCC
		20 54 420 495	GCC GCC GCC GCC
		20 54 420 495	GCC GCC GCC GCC
	seq.666	412 612	GCC GCG
		412 612	GCC GCG
10		412 612	GCC GCG
	seq.667	197 412	GCG GCC
		197 412	GCG GCC
		197 412	GCG GCC
15	seq.668	412 495 511	GCC GCC GCA
		412 495 511	GCC GCC GCA
		412 495 511	GCC GCC GCA
	seq.669	98 422	GCC GCC
		98 422	GCC GCC
		98 422	GCC GCC
20	seq.670	17 127 189	GCG GCT GCG
	seq.671	20 54 495	GCC GCC GCC
		20 54 495	GCC GCC GCC
		20 54 495	GCC GCC GCC
	seq.672	54 163	GCC GCT
25	seq.673	259 54	GCG GCC
		259 54	GCG GCC
		259 54	GCG GCC
	seq.674	335 399	GCT GCG
		335 399	GCT GCG
30		335 399	GCT GCG
	seq.675	221 432	GCA GCA
		221 432	GCA GCA
		221 432	GCA GCA
35	seq.676	259 516	GCG GCG
		259 516	GCG GCG
		259 516	GCG GCG
	seq.677	495 516	GCC GCG
		495 516	GCC GCG
		495 516	GCC GCG
40	seq.678	414 14	GCT GCC
		414 14	GCT GCC
		414 14	GCT GCC
	seq.679	74 402 495	GCG GCC GCC
		74 402 495	GCG GCC GCC
45		74 402 495	GCG GCC GCC
	seq.680	228 462 497	GCC GCC GCC
		228 462 497	GCC GCC GCC
		228 462 497	GCC GCC GCC
	seq.681	290 338	GCG GCC
50		290 338	GCG GCC
		290 338	GCG GCC

-74-

	seq.682	140 511	GCC GCA
		140 511	GCC GCA
		140 511	GCC GCA
5	seq.683	86 378	GCG GCG
		86 378	GCG GCG
		86 378	GCG GCG
	seq.684	54 86	GCC GCG
		54 86	GCC GCG
		54 86	GCC GCG
10	seq.685	214 495 140	GCG GCC GCC
		214 495 140	GCG GCC GCC
		214 495 140	GCG GCC GCC
	seq.686	495 511	GCC GCA
		495 511	GCC GCA
15		495 511	GCC GCA
	seq.687	495 54	GCC GCC
		495 54	GCC GCC
		495 54	GCC GCC
	seq.688	197 495	GCG GCC
20		197 495	GCG GCC
		197 495	GCG GCC
	seq.689	261 20	GCC GCC
		261 20	GCC GCC
		261 20	GCC GCC
25	seq.690	54 20	GCC GCC
	seq.691	197 420	GCG GCC
		197 420	GCG GCC
		197 420	GCG GCC
	seq.692	54 338 495	GCC GCC GCC
30		54 338 495	GCC GCC GCC
		54 338 495	GCC GCC GCC
	seq.693	197 427	GCG GCG
		197 427	GCG GCG
		197 427	GCG GCG
35	seq.694	54 228 370 387	GCC GCC GCC GCG
		54 228 370 387	GCC GCC GCC GCG
		54 228 370 387	GCC GCC GCC GCG
	seq.695	221 289	GCA GCC
		221 289	GCA GCC
40		221 289	GCA GCC
	seq.696	54 163	GCC GCT
		54 163	GCC GCT
	seq.697	341 407 420	GCC GCC GCC
		341 407 420	GCC GCC GCC
45		341 407 420	GCC GCC GCC
	seq.698	54 228	GCC GCC
		54 228	GCC GCC
		54 228	GCC GCC
	seq.699	96 125 511	GCA GCG GCA
50		96 125 511	GCA GCG GCA
		96 125 511	GCA GCG GCA

-75-

	seq.700	197 420	GCG GCC
		197 420	GCG GCC
		197 420	GCG GCC
5	seq.701	334 428 499	GCG GCT GCC
		334 428 499	GCG GCT GCC
		334 428 499	GCG GCT GCC
10	seq.702	197 414	GCG GCT
		197 414	GCG GCT
		197 414	GCG GCT
15	seq.703	30 54 127	GCG GCC GCT
		29 260	GCG GCG
20	seq.704	29 260	GCG GCG
		29 260	GCG GCG
		4 484	GCT GCC
25	seq.706	4 484	GCT GCC
		4 484	GCT GCC
		258 124 132	GCC GCC GCC
30	seq.707	258 124 132	GCC GCC GCC
		258 124 132	GCC GCC GCC
		231 497	GCC GCC
35	seq.708	231 497	GCC GCC
		231 497	GCC GCC
		221 258	GCA GCC
40	seq.709	221 258	GCA GCC
		221 258	GCA GCC
		234 264 326	GCG GCG GCC
45	seq.710	234 264 326	GCG GCG GCC
		234 264 326	GCG GCG GCC
		153 398	AGC GCG
50	seq.711	153 398	AGC GCG
		153 398	AGC GCG
		53 216	GCG GCC
55	seq.712	22 382	GCT GCG
		22 382	GCT GCG
		22 382	GCT GCG
60	seq.713	231 411	GCC GCA
		231 411	GCC GCA
		231 411	GCC GCA
65	seq.714	59 305	GCG GCC
		59 305	GCG GCC
		59 305	GCG GCC
70	seq.715	53 231	GCG GCC
		53 231	GCG GCC
		53 231	GCG GCC
75	seq.717	258 498	GCC GCT
		258 498	GCC GCT
		258 498	GCC GCT
80	seq.718	88 231	GCC GCC
		88 231	GCC GCC
		88 231	GCC GCC
85	seq.719	101 363	GCA GCC

-76-

	101 363	GCA GCC
	101 363	GCA GCC
	seq.720 354 132	GCC GCC
	354 132	GCC GCC
5	354 132	GCC GCC
	seq.726 598	GAC
	seq.727 598	AGC
	seq.728 600	CCG

- The above nucleic acid molecules are provided in plasmids, which
- 10 are introduced into cells to produce the encoded proteins. The analysis revealed the amino acid positions that affect Rep proteins activities. Changes of amino acids at any of the hit positions result in altered protein activity. Hit positions are numbered and referenced starting from amino acid 1 (nucleotide 321 in AAV-2 genome), also codon 1 of the protein
- 15 Rep78 coding sequence under control of p5 promoter of AAV-2: 4, 20, 22, 29, 32, 38, 39, 54, 59, 124, 125, 127, 132, 140, 161, 163, 193, 196, 197, 221, 228, 231, 234, 258, 260, 263, 264, 334, 335, 337, 342, 347, 350, 354, 363, 364, 367, 370, 376, 381, 389, 407, 411, 414, 420, 421, 422, 424, 428, 438, 440, 451, 460, 462, 484, 488,
- 20 495, 497, 498, 499, 503, 511, 512, 516, 517, 518, 542, 548, 598, 600 and 601. The encoded Rep78, Rep68, Rep 52 and Rep 40 proteins and rAAV encoding the mutant proteins are provided. The corresponding nucleic acid molecules, Rep proteins, rAAV and cells containing the nucleic acid molecules or rAAV in which the native proteins are from
- 25 other AAV serotypes, including, but are not limited to, AAV-1, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6.

Other hit positions identified include: 10, 64, 74, 86, 88, 101, 175, 237, 250, 334, 429 and 519.

- Also provided are nucleic acid molecules, the rAAV that encode
- 30 the mutant proteins, and the encoded proteins in which the native amino acid at each hit position is replaced with another amino acid, or is deleted, or contains additional amino acids at or adjacent to or near the

hit positions. In particular the following nucleic acid molecules and rAAV that encode proteins containing the following amino acid replacements or combinations thereof: T by N at Hit position 350; T by I at Hit position 462; P by R at Hit position 497; P by L at Hit position 497; P by Y at Hit 5 position 497; T by N at Hit position 517; L by S at hit position 542; R by S at hit position 548; G by D at Hit position 598; G by S at Hit position 598; V by P at Hit position 600; in order to increase Rep proteins activities in terms on AAV or rAAV productivity. The corresponding nucleic acid molecules, recombinant Rep proteins from the other 10 serotypes and the resulting rAAV are also provided (see Figs. 3 and the above Table for the corresponding position in AAV-1, AAV-3, AAV-3B, AAV-4, AAV-5 and AAV-6).

Mutant adeno-associated virus (AAV) Rep proteins and viruses encoding such proteins that include mutations at one or more of residues 15 64, 74, 88, 175, 237, 250 and 429, where residue 1 corresponds to residue 1 of the Rep78 protein encoded by nucleotides 321-323 of the AAV-2 genome, and where the amino acids are replaced as follows: L by A at position 64; P by A at position 74; Y by A at position 88; Y by A at position 175; T by A at position 237; T by A at position 250; D by A at 20 position 429 are provided. Nucleic acid molecules encoding these viruses and the mutant proteins are also provided.

Also provided are nucleic acid molecules produced from any of the above-noted nucleic acid molecules by any directed evolution method, including, but are not limited to, re-synthesis, mutagenesis, recombination 25 and gene shuffling and any way by combining any combination of the molecules, *i.e.*, one, two by one, two by two, .....n by n, where n is the number of molecules to be combined (*i.e.*, combining all together). The resulting recombinant AAV and encoded proteins are also provided.

-78-

Also provided are nucleic acid molecule in which additional amino acids surrounding each hit, such as one, two, three . . . ten or more, amino acids are systematically replaced, such that the resulting Rep protein(s) has increased or decreased activity. Increased activity as

- 5 assessed by increased recombinant virus production in suitable cells is of particular interest for production of recombinant viruses for use, for example, in gene therapy.

Also provided are combinations of the above noted mutants in which several of the noted amino acids are changed and optionally

- 10 additional amino acids surrounding each hit, such as one, two, three . . . ten or more, are replaced.

For all of the mutant proteins provided herein those with increased activity, such as an increase in titer of rAAV when virus containing such mutations and/or expressing such mutant proteins are replicated, are of 15 particular interest. Such mutations and proteins are provided herein and may be made by the methods herein, including by combining any of the mutations provided herein to produce additional mutant proteins that have altered biological activity, particularly increased activity, compared to the wild-type.

- 20 The nucleic acid molecules of SEQ ID Nos. 563-725 and the encoded proteins (SEQ ID Nos. 1-562 and 726-728) are also provided. Recombinant AAV and cells containing the encoding nucleic acids are provided, as are the AAV produced upon replication of the AAV in the cells.

- 25 Methods of *in vivo* or *in vitro* production of AAV or rAAV using any of the above nucleic acid molecules or cells for intracellular expression of rep proteins or the rep gene mutants are provided. *In vitro* production is effected using cell free systems, expression or replication

-79-

and/or virus assembly. *In vivo* production is effected in mammalian cells that also contain any requisite *cis* acting elements required for packaging.

Also provided are nucleic acid molecules and rAAV (any serotype) in which position 630 (or the corresponding position in another serotype; 5 see Figs. 3 and the table above) has been changed. Changes at this position and the region around it lead to changes in the activity or in the quantities of the Rep or Cap proteins and/or the amount of AAV or rAAV produced in cells transduced with AAV encoding such mutants. Such mutations include tgc to gcg change (SEQ ID No. 721). Mutations at any 10 position surrounding the codon position 630 that increase or decrease the Rep or Cap proteins quantities or activities are also provided. Methods using the rAAV (any serotype) that contain nucleic acid molecules with a mutation at position 630 or within 1, 2, 3 . . . 10 or more bases thereof for the intracellular expression rep proteins or the rep gene mutants 15 covered by claims 10 to 13, for the production of AAV or rAAV (either *in vitro*, *in vivo* or *ex vivo*) are provided. *In vitro* methods include cell free systems, expression or replication and/or virus assembly.

Also provided are rAAV (and other serotypes with corresponding changes) and nucleic acid molecules encoding an amino acid replacement 20 by N at Hit position 350 of AAV- 1, AAV-3, AAV-3B, AAV-4 and AAV-6 or at Hit position 346 of AAV-5; by I at Hit position 462 of AAV-1, AAV-3, AAV-3B, AAV-4 and AAV-6 or at Hit position 458 of AAV-5; by either R, L or Y at Hit position 497 of AAV-1, AAV-3, AAV-3B, AAV-4 and AAV-6 or at Hit position 493 of AAV-5; by N at Hit position 517 of 25 AAV-1, AAV-3, AAV-3B, AAV-4 and AAV-6 or at Hit position 535 of AAV-5; by S at hit position 543 of AAV-1 and AAV-6 or at hit position 542 of AAV-3, AAV-3B and AAV-4 or at hit position 561 of AAV-5; by S at hit position 549 of AAV-1 and AAV-6 or at hit position 548 of AAV-3, AAV-3B and AAV-4 or at hit position 567 of AAV-5; by either D or S at

-80-

Hit position 599 of AAV-1, AAV-4 and AAV-6 or at Hit position 600 of  
AAV-3 and AAV-3B; by P at Hit position 602 of AAV-1, AAV-4 and AAV-  
6 or at hit position 603 of AAV-3 and AAV-3B or at hit position 589 of  
AAV-5 in order to increase Rep proteins activities as assessed by AAV or

- 5 rAAV productivity. Methods using such AAV for expression of the  
encoded proteins and production of AAV are also provided.

Since modifications will be apparent to those of skill in this art, it is  
intended that this invention be limited only by the scope of the appended  
10 claims.

-81-

**WHAT IS CLAIMED IS:**

1. An adeno-associated virus (AAV), comprising nucleic acid encoding the sequence of amino acids in any of SEQ ID Nos. 1-562 and 726-728 or encoding a sequence of amino acids encoded by SEQ ID Nos. 5 722-725.
2. The AAV of claim 1, wherein the sequence of nucleotides encoding the sequence of amino acids is set forth in SEQ ID Nos. 563-725.
3. The AAV of claim 1 that has an altered activity in a Rep 10 protein and/or a capsid protein.
4. The AAV of claim 3, wherein the alteration leads to greater activity in the Rep gene manifested as an increased titer of virus upon introduction and replication in a host cell compared to the titer of virus upon introduction and replication of a wild type Rep gene.
- 15 5. The AAV of claim 1 that is of serotype AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 or AAV-6.
6. A mutant adeno-associate virus (AAV) Rep protein, comprising mutations at one or more of residues 4, 20, 22, 29, 32, 38, 39, 54, 59, 124, 125, 127, 132, 140, 161, 163, 193, 196, 197, 221, 20 228, 231, 234, 258, 260, 263, 264, 334, 335, 337, 342, 347, 350, 354, 363, 364, 367, 370, 376, 381, 389, 407, 411, 414, 420, 421, 422, 424, 428, 438, 440, 451, 460, 462, 484, 488, 495, 497, 498, 499, 503, 511, 512, 516, 517, 518, 542, 548, 598, 600 and 601 of 25 AAV-2 or the corresponding residues in other serotypes, wherein residue 1 corresponds to residue 1 of the Rep78 protein encoded by nucleotides 321-323 of the AAV-2 genome, wherein the mutations comprise insertions, deletions or replacements of the native amino acid residue(s).
7. The Rep protein of claim 6 that is Rep 78, Rep 68, Rep 52 or Rep 40.

-82-

8. The mutant AAV Rep protein of claim 6, wherein the AAV is an AAV-1, AAV-2, AAV-3, AAV-3b, AAV-4, AAV-5 or AAV-6, wherein the mutation is in the equivalent position in each serotype, wherein the listed residues are the positions in AAV-2.

5 9. A mutant AAV Rep protein of claim 6 that has increased activity compared to the native protein, wherein activity is assessed by measuring viral production when an AAV that encodes the protein is introduced into a cell under conditions wherein the virus replicates.

10 10. A mutant AAV Rep protein of claim 6 that has decreased activity compared to the native protein, wherein activity is assessed by measuring viral production when an AAV that encodes the protein is introduced into a cell under conditions wherein the virus replicates.

11. A mutant Rep protein of claim 6, further comprising a mutation at one or more of residues 10, 64, 74, 86, 88, 101, 175, 237, 15 250, 334, 429 and 519.

12. The mutant Rep protein of claim 6, wherein the amino acids are replaced as follows: T by N at position 350; T by I at position 462; P by R at position 497; P by L at position 497; P by Y at position 497; T by N at position 517; G by D at position 598; G by S at position 598; V by P 20 at position 600, whereby the activity of the Rep protein is increased as assessed by rAAV production compared to the native Rep protein.

13. A mutant Rep protein of claim 6, comprising two or more of the mutations.

14. A mutant adeno-associate virus (AAV) Rep protein, 25 comprising mutations at one or more of residues 64, 74, 88, 175, 237, 250 and 429, wherein: residue 1 corresponds to residue 1 of the Rep78 protein encoding by nucleotides 321-323 of the AAV-2 genome; wherein the amino acids are replaced as follows:

L by A at position 64;

-83-

P by A at position 74;  
Y by A at position 88;  
Y by A at position 175;  
T by A at position 237;  
5 T by A at position 250;  
D by A at position 429;

the mutations comprise insertions, deletions or replacements of the native amino acid residue.

15. A nucleic acid molecule encoding the protein of claim 6.
- 10 16. A recombinant AAV comprising the nucleic acid molecule of claim 15.
17. A eukaryotic cell, comprising the recombinant AAV of claim 16.
18. A collection of nucleic acid molecules comprising a plurality of 15 the molecules of claim 17.
19. A collection of nucleic acid molecules comprising a plurality of the molecules of claim 15.
20. An isolated nucleic acid molecule encoding the proteins of SEQ ID Nos. 1-562 and 726-728 or encoding a sequence of amino acids encoded by SEQ ID Nos. 722-725.
21. A Rep protein of any of SEQ ID Nos. 1-562 and 726-728 or encoding a sequence of amino acids encoded by SEQ ID Nos. 722-725.
22. A Rep protein encoded by any of SEQ ID Nos. 564-725.
23. A method for intracellular expression of a mutant Rep 25 protein, comprising:
  - introducing the recombinant AAV of claim 16 into a host cell; and
  - culturing the cell, under conditions and in which the AAV Rep proteins are expressed.

-84-

24. The method of claim 23, wherein the AAV replicates.

25. An AAV genome, comprising a mutation at one or more of nucleotides corresponding to nucleotides 2209-2211 of the AAV-2 genome, which encode amino acid residue 630 of the Rep78 protein,

5 wherein:

the mutation is a deletion, insertion or replacement of a nucleotide; and the mutation results in a change in the activity or in the quantities of the Rep or Cap proteins as assessed by the level of replication of the AAV genome.

10 26. The AAV genome of claim 25, wherein the mutation at position 630 is a tgc to gcg and the intron comprises the sequence (SEQ ID No. 722):

gtacccaaacaaatgttctcgtcacgtggcatgaatctgatgctgtttccctgc  
agacaatgcgagagaatgaatcagaattcaaataatctgcttcactcacggacaga

15 aagactgttagagtgtttccctgtcagaatctcaacccgtttctgtcgtaa  
aaaggcgtatcagaaactgtgctacattcatatatcatggaaaagggtgccagac  
gcttgactgcctgcgtatctggtaatgtggattggatgactgcacatcttgaac  
aataaatgattaaatcaggtatggcgcgatggttatcttccag.

27. The AAV genome of claim 25, wherein the mutation at  
20 position 630 is a tgc to cgc and the intron comprises the sequence (SEQ ID No. 723):

gtacccaaacaaatgttctcgtcacgtggcatgaatctgatgctgtttcc  
ctgcagacaatgcgagagaatgaatcagaattcaaataatctgcttcactcac  
ggacagaaaagactgttagagtgtttccctgtcagaatctcaacccgtttc  
25 tgtcgtaaaaaggcgatcagaaactgtgctacattcatatatcatggaaa  
agggtgccagacgcgttgcactgcctgcgtatctggtaatgtggattggatgac  
tgcatcttgaacaataatgattaaatcaggtatggccgcgtatggttatc  
ttccag.

-85-

28. The AAV genome of claim 25, wherein the mutation at position 630 is a tgc to cct and the intron comprises the sequence (SEQ ID No. 724):

gtacaaaaacaatgttctcgtcacgtggcatgaatctgatgctgttt  
5 ccctgcagacaatgcgagagaatgaatcagaattcaaatatctgcttcac  
tcacggacagaaagactgttagagtgccttccgtcagaatctcaac  
ccgttctgtcgtaaaaaggcgtatcagaaactgtgctacattcatcat  
atcatggaaagggccagacgcctgcactgcctgcgatctggtaatgt  
ggatttggatgactgcatttgaacaataatgattaaatcaggt  
10 atggccctcgatggttatcttccag.

29. The AAV genome of claim 25, wherein the mutation at position 630 is a tgc to tca and the intron comprises the sequence (SEQ ID No. 725):

gtacaaaaacaatgttctcgtcacgtggcatgaatctgatgctgtttcc  
15 ctgcagacaatgcgagagaatgaatcagaattcaaatatctgcttcactca  
cgacagaaagactgttagagtgccttccgtcagaatctcaacccgt  
ttctgtcgtaaaaaggcgtatcagaaactgtgctacattcatcatatcat  
ggaaagggccagacgcctgcactgcctgcgatctggtaatgtggattt  
ggatgactgcatttgaacaataatgattaaatcaggtatggctacg  
20 atggttatcttccag.

30. A method for intracellular expression of a mutant Rep protein, comprising:

introducing the recombinant AAV of claim 25 into a host cell; and

25 culturing the cell, under conditions and in which the AAV Rep proteins and/or cap proteins are expressed.

31. The method of claim 30, wherein the AAV replicates.

32. The AAV genome of claim 25, wherein the AAV is of serotype AAV-1, AAV-3, AAV-3B, AAV-4, AAV-5 or AAV-6.

-86-

33. A method of titrating virus by a method designated tagged replication and expression enhancement, comprising:

- (i) incubating host cells with a reporter virus vector and with a titrating virus of unknown titer, wherein a titrating virus  
5 increases or decreases the output signal from the reporter virus; and
- (ii) measuring the output signal of the reporter virus and determining the titer of the reporter virus; and
- (ii) determining the titer of the titrating virus by comparing the titer of the reporter virus in the presence and absence of  
10 the titrating virus.

34. A process for the production of an adeno-associated virus (AAV) protein or a recombinant AAV having a predetermined property, comprising:

- (a) producing a population of sets of nucleic acid molecules that  
15 encode modified forms of a target protein;
- (b) introducing each set of nucleic acid molecules into host cells and expressing the encoded protein, wherein the host cells are present in an addressable array;
- (c) individually screening the sets of encoded proteins to identify  
20 one or more proteins that have activity that differs from the target protein, wherein each such protein is designated a hit;
- (d) modifying the nucleic acid molecules that encode the hits, to produce a set of nucleic acid molecules that encode modified hits, wherein the nucleic acid molecules comprise rAAV vectors;
- 25 (e) introducing the each set of nucleic acids that encode the modified hits into cells; and
- (f) individually screening the sets of cells that contain the nucleic acid molecules that encode the modified hits to identify one or more cells that encodes a protein that has activity that differs from the target protein

-87-

and has properties that differ from the original hits, wherein each such protein is designated a lead.

35. The process of claim 34, wherein the cells are eukaryotic cells that are transduced with the vectors.

5 36. The method of claim 35, wherein at step (f) the titer of the viral vectors in each set of cells is determined.

37. The method of claim 36, wherein the target protein is a protein involved in viral replication.

10 38. The method of claim 37, wherein the target protein is a Rep protein.

39. The AAV mutant Rep protein of claim 6 that binds to a sequence from a papillomavirus, oncogene or human immunodeficiency virus (HIV) with different affinity from a wild-type AAV Rep protein.

15 40. A fusion protein, comprising the *tat* protein of HIV and the mutant Rep protein of claim 39.

41. The fusion protein of claim 40, wherein the HIV is HIV-1.

42. A pharmaceutical composition, comprising the protein of claim 39 in a pharmaceutically acceptable carrier.

20 43. A recombinant adeno-associated virus (rAAV) that encodes a mutant Rep protein that has increased activity, wherein increased activity of a Rep protein is manifested as an increased titer of virus upon introduction and replication in a host cell compared to the titer of virus upon introduction and replication of a wild type Rep gene.

44. A mutant AAV Rep protein that has increased activity,  
25 wherein increased activity of a Rep protein is manifested as an increased titer of virus upon introduction and replication in a host cell compared to the titer of virus upon introduction and replication of a wild type Rep gene.

-88-

45. A nucleic acid molecule that encodes that mutant Rep protein of claim 44.
46. A cell, comprising the nucleic acid molecule of claim 45.
47. A rAAV, comprising the nucleic acid molecule of claim 45.
- 5 48. A cell, comprising the rAAV of claim 47.
49. A method for production of rAAV, comprising:  
introducing the rAAV of claim 47 into a cell under conditions  
whereby the virus replicates to produce encapsulated rAAV.
- 10 50. A method for the production of mutant Rep protein comprising  
expressing the nucleic acid molecule of claim 45.
51. The method of claim 50, wherein expression is effected *in vivo*.
52. The method of claim 50, wherein expression is effected *in vitro*.
- 15 53. A method for producing Rep protein in a host cell, comprising:  
expressing the protein encoded by the nucleic acid encoding  
the protein of claim 44, wherein the method is performed *in vitro* or *in vivo*.
54. The method of claim 53, wherein the nucleic acid is  
20 introduced into a cell.
55. The method of claim 53, wherein expression is effected in a cell-free system.
56. A method of treating or inhibiting infection by human papilloma virus or a human immunodeficiency virus, comprising  
25 administering, to a subject exposed to the virus or infected with the virus,  
a composition containing a rAAV of claim 47.
57. A nucleic acid molecule encoding the protein of claim 7.
58. A nucleic acid molecule encoding the protein of claim 8.
59. A nucleic acid molecule encoding the protein of claim 9.

-89-

60. A nucleic acid molecule encoding the protein of claim 10.
61. A nucleic acid molecule encoding the protein of claim 11.
62. A nucleic acid molecule encoding the protein of claim 12.
63. A nucleic acid molecule encoding the protein of claim 13.
- 5 64. A nucleic acid molecule encoding the protein of claim 14.
65. A recombinant AAV comprising the nucleic acid molecule of  
claim 57.
66. A recombinant AAV comprising the nucleic acid molecule of  
claim 58.
- 10 67. A recombinant AAV comprising the nucleic acid molecule of  
claim 59.
68. A recombinant AAV comprising the nucleic acid molecule of  
claim 60.
69. A recombinant AAV comprising the nucleic acid molecule of  
15 claim 61.
70. A recombinant AAV comprising the nucleic acid molecule of  
claim 62.
71. A recombinant AAV comprising the nucleic acid molecule of  
claim 63.
- 20 72. A recombinant AAV comprising the nucleic acid molecule of  
claim 64.
73. A cell, comprising the recombinant AAV of claim 65.
74. A cell, comprising the recombinant AAV of claim 66.
75. A cell, comprising the recombinant AAV of claim 67.
- 25 76. A cell, comprising the recombinant AAV of claim 68.
77. A cell, comprising the recombinant AAV of claim 69.
78. A cell, comprising the recombinant AAV of claim 70.
79. A cell, comprising the recombinant AAV of claim 71.
80. A cell, comprising the recombinant AAV of claim 72.

-90-

81. A method for intracellular expression of a mutant Rep protein, comprising:

culturing the cell of claim 73 under conditions and in which the AAV Rep proteins are expressed.

5 82. The method of claim 81, wherein the AAV replicate.

83. A method for intracellular expression of a mutant Rep protein, comprising culturing the cell of claim 74 under conditions in which the AAV Rep proteins are expressed.

84. The method of claim 83, wherein the AAV replicate.

10 85. A method of altering expression of a gene, comprising contacting the gene with a mutant rep protein that has increased activity, wherein increased activity of a Rep protein is manifested as an increased titer of virus upon introduction and replication in a host cell compared to the titer of virus upon introduction and replication of a wild type Rep 15 gene.

86. The method of claim 85, wherein the gene is a viral gene.

87. The method of claim 85, wherein the gene is a cellular gene.

88. The mutant protein of claim 6, wherein serotype is AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 or AAV-6.

20 89. The protein of claim 44, wherein the mutation is at a residue corresponding to one or more of residues 350, 462, 497, 517, 542, 548, 598, 600 and 630 of AAV-2.

90. The mutant protein of claim 89, wherein serotype is AAV-1, AAV-2, AAV-3, AAV-3B, AAV-4, AAV-5 or AAV-6.

25 91. The AAV mutant Rep protein of claim 44 that binds to a sequence from a papillomavirus, oncogene or human immunodeficiency virus (HIV) with different affinity from a wild-type AAV Rep protein.

92. A pharmaceutical composition, comprising the protein of claim 91 in a pharmaceutically acceptable carrier.

-91-

93. A pharmaceutical composition, comprising the rAAV of claim 47 in a pharmaceutically acceptable carrier.

1/4

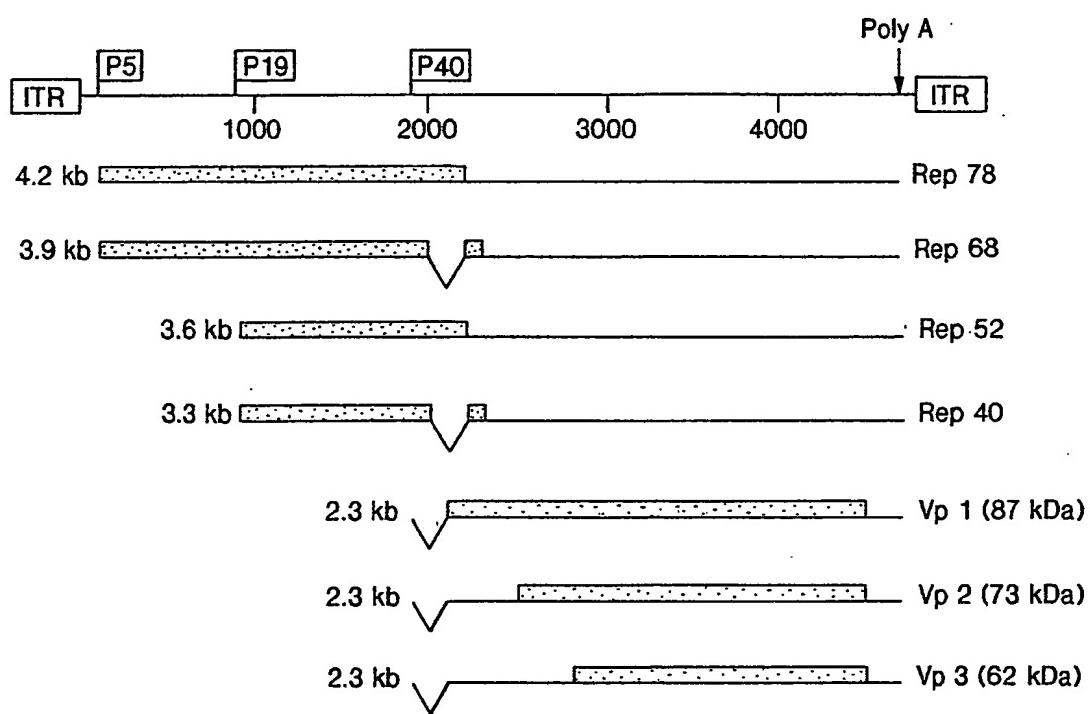


FIG. 1

2/4

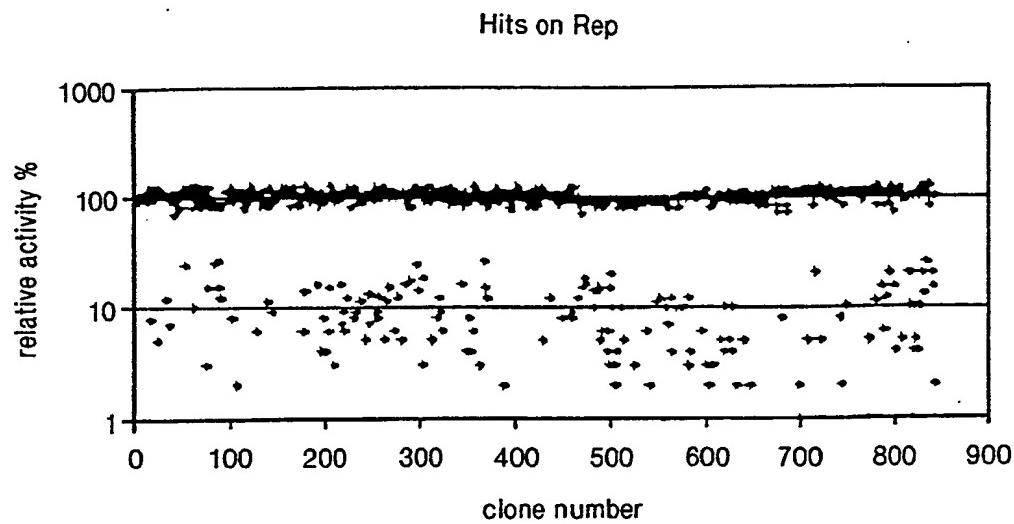


FIG. 2A

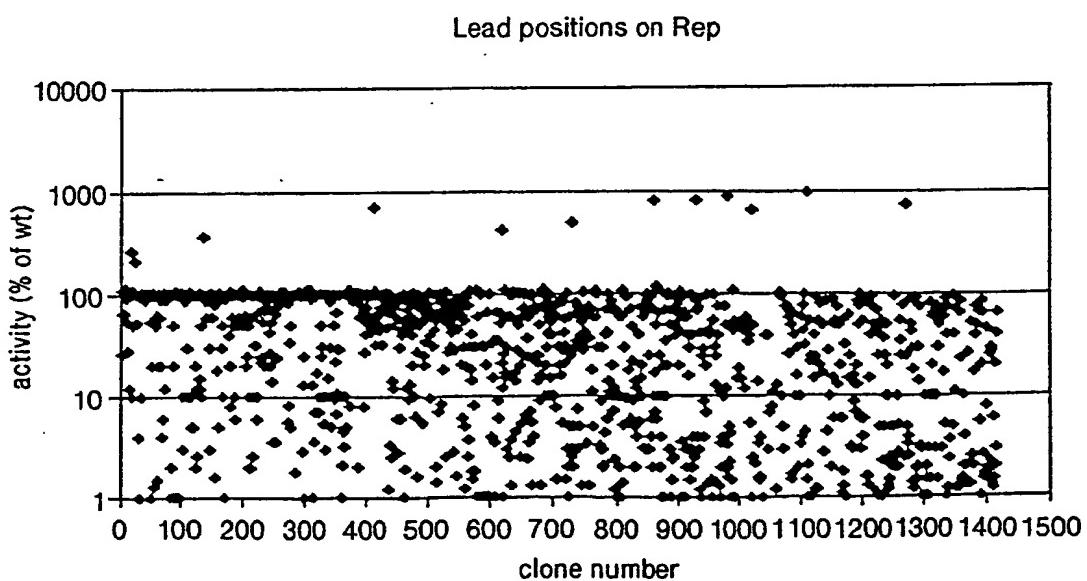


FIG. 2B

3/4

	10	20	30	40	50	60	
1	MPGFYEIVLKVPSDLDEHLPGISDSFVSWVAEKEWELPPSDMDLNLIEQAPLTVAEKLQ-						60
2	MPGFYEIVLKVPSDLDEHLPGISDSFVNWVAEKEWELPPSDMDLNLIEQAPLTVAEKLQ						60
3	MPGFYEIVLKVPSDLDEHLPGISNSFVNWVAEKEWELPPSDMDPNLIEQAPLTVAEKLQ						60
4	MPGFYEIVLKVPSDLDEHLPGISNSFVNWVAEKEWELPPSDMDPNLIEQAPLTVAEKLQ						60
5	MPGFYEIVLKVPSDLDEHLPGISDSFVSWVAEKEWELPPSDMDLNLIEQAPLTVAEKLQ						60
6	MPGFYEIVLKVPSDLGHLPGISDSFVNWVAEKEWELPPSDMDLNLIEQAPLTVAEKLQ						60
7	MATFYEVIVRVPFDVEEHLPGISDSFVDWVTGQIWELPPESDLNLTLVEQPQLTVADRIR						60
C	M**FYE** : *VP*D***HLPGIS+SFV:WV****WELPP*SD***+*L*EQ**LTVA****						
	70	80	90	100	110	120	
1	RDFLVQWRRVSKAPEALFFVQFEKGESYFHLHILVETTVGKSMVLGRFLSQIRDKLVQTI						120
2	RDFLVQWRRVSKAPEALFFVQFEKGESYFHLHILVETTVGKSMVLGRFLSQIRDKLVQTI						120
3	REFLVEWRRVSKAPEALFFVQFEKGESYFHLHVLIETIGVKSMVVGRYVSQIKEKLVTRI						120
4	REFLVEWRRVSKAPEALFFVQFEKGESYFHLHVLIETIGVKSMVVGRYVSQIKEKLVTRI						120
5	REFLVEWRRVSKAPEALFFVQFEKGESYFHLHVLIETIGVKSMVVGRYVSQIKEKLVTRI						120
6	RDLTEWRRVSKAPEALFFVQFEKGESYFHMHVLTETGVKSMVLGRFLSQIREKLIQRI						120
7	RVFLYEWNKFSQ-E SKFFVQFEKGSEYFHLHTLVTSGISMVGLGRYVSQIRAQLVKVV						119
C	R : FL++W***SK**E***FFVQFEKG+:YFH*H:L+ET:G**SMV:GR::SQI:::L*:::*						
	130	140	150	160	170	180	
1	YRGIEPTLPNWFAVTKTRNGAGGNKVVDECYIPNYLLPKTQPELQWAUTNMEEYISACL						180
2	YRGIEPTLPNWFAVTKTRNGAGGNKVVDECYIPNYLLPKTQPELQWAUTNMEEYISACL						180
3	YRGVEPQLPNWFAVTKTRNGAGGNKVVDDCYIPNYLLPKTQPELQWAUTNMMDQYLSACL						180
4	YRGVEPQLPNWFAVTKTRNGAGGNKVVDDCYIPNYLLPKTQPELQWAUTNMMDQYLSACL						180
5	YRGVEPQLPNWFAVTKTRNGAGGNKVVDDCYIPNYLLPKTQPELQWAUTNMMDQYISACL						180
6	YRGIEPTLPNWFAVTKTRNGAGGNKVVDECYIPNYLLPKTQPELQWAUTNMEQYLSACL						180
7	FQGIEPQINDWVAITKVKK--GGANKVVDGSGYIPAYLLPKVQPELQWAUTNLDEYKLAAL						177
C	**G:EP:***W*A*TK*****GG*NVVD:YIP*YLLPK*QPELQWAUTN*:Y: A*L						
	190	200	210	220	230	240	
1	NLAERKRLVAQHLTHVSQTQEQNKENLNPNNSDAPVIRSKTSARYMELVGWLVRGITSEK						240
2	NLAERKRLVAHDLTHVSQTQEQNKENLNPNNSDAPVIRSKTSARYMELVGWLVRGITSEK						240
3	NLAERKRLVAQHLTHVSQTQEQNKENQNPNNSDAPVIRSKTSARYMELVGWLVRGITSEK						240
4	NLAERKRLVAQHLTHVSQTQEQNKENQNPNNSDAPVIRSKTSARYMELVGWLVRGITSEK						240
5	NLAERKRLVAQHLTHVSQTQEQNKENQNPNNSDAPVIRSKTSARYMELVGWLVRGITSEK						240
6	NLTERKRLVAQHLTHVSQTQEQNKENQNPNNSDAPVIRSKTSARYMELVGWLVDKGITSEK						240
7	NLEERKRLVAQFLAESSQRS-QEAASQREFSADPVIKSKTSQLYMAVLNWLVHGITSEK						236
C	NL+ERKRLVA*+L***SQ***Q*****S***PVI*SKTS**YM*LV*WLV*+GITSEK						
	250	260	270	280	290	300	
1	QWIQEDQASYISFNAA NSRSQIKAALDNAGKIMALT KSA PDYL VGPAPPADIKTNRIYR						300
2	QWIQEDQASYISFNAA NSRSQIKAALDNAGKIMALT KSA PDYL VGPAPPADIKTNRIYR						300
3	QWIQEDQASYISFNAA NSRSQIKAALDNASKIMSLTKTAPDYL VGSNP PEDITKNRIYQ						300
4	QWIQEDQASYISFNAA NSRSQIKAALDNASKIMSLTKTAPDYL VGSNP PEDITKNRIYQ						300
5	QWIQEDQASYISFNAA NSRSQIKAALDNASKIMSLTKTAPDYL VGSNP PEDITKNRIYR						300
6	QWIQEDQASYISFNAA NSRSQIKAALDNAGKIMSLTKTAPDYL VGQQPVEDISSNRIYK						300
7	QWIQENQESYLSFNSTGNSRSQIKAALDNATKIMSLTKSAVDYL VGSSVPEDISKNRIWQ						296
C	QWIQE*Q*SY*SFN***NSRSQIKAALDNA:KIM+LTK:A*DYLVG:***+DI: :NRI*:						
	310	320	330	340	350	360	
1	I LE L NGYEPAYAGSVFLGWAQKRFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
2	I LE L NGYDPAYAGSVFLGWAQKRFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
3	I LE L NGYDPQYAASVFLGWAQKKFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
4	I LE L NGYDPQYAASVFLGWAQKKFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
5	I LE M NGYDPQYAASVFLGWAQKKFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
6	I LE L NGYDPQYAASVFLGWAQKKFGKRNTIWLFGPATTGKTNIAEAIAAHAVPFYGCVNWT						360
7	I FEMNGYDPAYAGSILYGCWCQRSFNKRNTVWLYGPATTGKTNIAEAIAHTVPFYGCVNWT						356
C	I *E+NGY*P:YA:S***GW***:F*KRNT*WL*GPATTGKTNIAEAIAH+VPFYGCVNWT						

FIG. 3A

4/4

	370	380	390	400	410	420	
1	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTS						420
2	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTS						420
3	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIEPTPVIVTS						420
4	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIEPTPVIVTS						420
5	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTS						420
6	NENFPFNDCVDKMVIWWEEGKMTAKVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTS						420
7	NENFPFNDCVDKMLIWWEEGKMTNKVVESAKAILGGSKVRVDQKCKSSVQIDSTPVIVTS						416
C	NENFPFNDCVDKM*IWWEEGKMT*KVVESAKAILGGSKVRVDQKCKSS*QI+*TPVIVTS						
	430	440	450	460	470	480	
1	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLHDFGKVTQEVKEFFRWAQDHVTEV						480
2	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLHDFGKVTQEVKEFFRWAQDHVTEV						480
3	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLDHDFGKVTQEVKDFFRWASDHVTDV						480
4	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLDHDFGKVTQEVKDFFRWASDHVTDV						480
5	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLDHDFGKVTQEVKDFFRWASDHVTDV						480
6	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLDHDFGKVTQEVKDFFRWASDHVVEV						480
7	NTNMCAVIDGNSTTFEHQQLQDRMFKFELTRRLDHDFGKVTQEVKDFFRWASDHVVEV						476
C	NTNMC*V*DGNSTTFEHQQLPLQDRMFKFELTKRLPPDFGKTFKQEVKDFFAWAKVNQPV						
	NTNMC*V*DGNSTTFEHQQLPLQDRMFKFELT+RL:DFGK*TKQEVK+FF*WA:****+:V						
	490	500			510	520	
1	AHEFYVRKGANKRPAAPDDADKSEPKRA-----				-CPSVADPSTS DAE		522
2	AHEFYVRKGANKRPAAPDDADKSEPKRA-----				-CPSVADPSTS DAE		522
3	AHEFYVRKGAKKRPA SNDADVSEPKRQ-----				-CTSLAQPTTSDAEA		522
4	AHEFYVRKGAKKRPA SNDADVSEPKRQ-----				-CTSLAQPTTSDAEA		522
5	THEFYVRKGARKRPA PNDADISEPKRA-----				-CPSVAQPSTS DAE		522
6	EHEFYVKGGAKKRPA PSDADISEPKRV-----				-RESVAQPSTS DAE		522
7	THEFKVPRELAGTKGAEKSLKRPLGDVTNTSYKSLEKRARLSFVPETPRSSDVTVDPAPL						536
C	:HEF*V+**A:***A:***.*****:				+*: *: *: *: *: *: *: A*:		
	530	540	550	560	570	580	
1	APVDFADRYQNKC SRAGMLQMLFPCKTCERMNQFNICFTHGTRDCSEC FP--GVSESQ						580
2	APVDFADRYQNKC SRAGMLQMLFPCKTCERMNQFNICFTHGTRDCSEC FP--GVSESQ						580
3	P-ADYADRYQNKC SRHVGMLMLFPCKTCERMNQISNVCFTHQRDCGECPGMSES QPV						581
4	P-ADYADRYQNKC SRHVGMLMLFPCKTCERMNQISNVCFTHQRDCGECPGMSES QPV						581
5	P-VDYADRYQNKC SRHVGMLMLFPCRQCERMNQVDICFTHGVMDCAE CF P-VSES QPV						580
6	S-INYADRYQNKC SRHVGMLMLFPCRQCERMNQNSNICFTHGQKD CLE CF P--VSES QPV						579
7	RPLNWNSRYDKCDYHAQFDNISNKDECEYLNRGKNGCICHNVTHCQICHG-----						588
C	:+: +: **RY**KC**H:***: ***C: :CE**N: :*: C**H: :*C.*C**. .: +: :						
	590	600	610	620			
1	PVVRKRTYRKLCIAIHLLGRAPEIAC SACDLVNVLDLDCVSEQ						623
2	PVVRKRTYRKLCIAIHLLGRAPEIAC SACDLVNVLDLDCVSEQ						623
3	SVVKKKTYQKLCP IHHILGRAPEIAC SACDLANVLDLDCVSEQ						624
4	SVVKKKTYQKLCP IHHILGRAPEIAC SACDLANVLDLDCVSEQ						624
5	SVVRKRTYQKLCP IHHIMGRAPEVAC SACELANVLDLDCDMEQ						623
6	VSVVKKAYQKLCP IHHIMGRAPEVAC SACELANVLDLDCDMEQ						621
7	-----IPPW EKENLSDFGDFDDANKEQ						610
C	:+: *: +*: ***: ***: : +: +: +: *: *: +: +: *: *: D*DD*: :EQ						

FIG. 3B

-1-

SEQUENCE LISTING

<110> Vega, Manuel  
Drittanti, Lila  
Flaux, Marjorie

<120> MUTANT RECOMBINANT ADENO-ASSOCIATED VIRUSES

<130> 37851-912PC

<140> To Be Assigned  
<141> Herewith

<150> 60/315,382  
<151> 2001-08-27

<160> 735

<170> FastSEQ for Windows Version 4.0

<210> 1  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 4 GCT

<400> 1  
Thr Ala Gly Ala Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255

-2-

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 2  
 <211> 536  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep68 4 GCT

<400> 2  
 Thr Ala Gly Ala Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile

-3-

Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
35								40							45
50								55							60
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65								70							80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
								85							95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
								100							110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
								115							125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
								130							140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
								145							160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
								165							175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
								180							190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
								195							205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210							220
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
								225							240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								245							255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
								260							270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
								275							285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Glu	Leu	
								290							300
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
								305							320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
								325							335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
								340							350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
								355							365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
								370							380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
								385							400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
								405							415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
								420							430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
								435							445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
								450							460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
								465							480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
								485							495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
								500							510
Ala	Gln	Pro	Ser	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp	
								515							525

-4-

Arg Leu Ala Arg Gly His Ser Leu  
530 535

<210> 3  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 10 GCG

<400> 3  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Ala Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg

-5-

385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
565	570	575	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

&lt;210&gt; 4

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 10 GCG

&lt;400&gt; 4

Thr Ala Gly Phe Tyr Glu Ile Val Ile Ala Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	

-6-

Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180                       185                       190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195                       200                       205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
     210                       215                       220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     225                       230                       235                       240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     245                       250                       255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     260                       265                       270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     275                       280                       285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     290                       295                       300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     305                       310                       315                       320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     325                       330                       335  
 Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     340                       345                       350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     355                       360                       365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     370                       375                       380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     385                       390                       395                       400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     405                       410                       415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     420                       425                       430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435                       440                       445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450                       455                       460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465                       470                       475                       480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485                       490                       495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     500                       505                       510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     515                       520                       525  
 Arg Leu Ala Arg Gly His Ser Leu  
     530                       535

<210> 5  
 <211> 621  
 <212> PRT

<213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep78 20 GCC

<400> 5  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1               5                       10                       15  
 Glu His Leu Ala Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                       25                       30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile

35	40	45													
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
50								55		60					
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65								70		75					80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
								85		90					95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
								100		105					110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Ile
								115		120					125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
								130		135					140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
								145		150					160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
								165		170					175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
								180		185					190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
								195		200					205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210		215					220
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
								225		230					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								245		250					255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
								260		265					270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
								275		280					285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
								290		295					300
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
								305		310					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
								325		330					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
								340		345					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
								355		360					365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
								370		375					380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Ser	Lys	Val	Arg	
								385		390					400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
								405		410					415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
								420		425					430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
								435		440					445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
								450		455					460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
								465		470					480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
								485		490					495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
								500		505					510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
								515		520					525

-8-

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 6  
 <211> 536  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep 68 20 GCC

<400> 6  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Ala Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala

-9-

305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		

&lt;210&gt; 7

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 22 GCT

&lt;400&gt; 7

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ala Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	

-10-

Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
180							185							190	
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
195							200							205	
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
210							215							220	
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225							230							235	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							245							250	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
							260							265	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
							275							280	
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
							290							295	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
							305							310	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325							330	
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
							340							345	
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355							360	
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370							375	
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
							385							390	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405							410	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420							425	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435							440	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450							455	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
							465							470	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485							490	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500							505	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515							520	
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
							530							535	
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
							545							550	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
							565							570	
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
							580							585	
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595							600	
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			605
							610							615	
														620	

<210> 8  
<211> 536  
<212> PRT  
<213> Artificial Sequence

-11-

<220>  
 <223> Mutant rep protein: rep68 22 GCT  
 <400> 8  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ala Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln

-12-

450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 9

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 29 GCG

<400> 9	10	15
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Ala Val Ala Glu		
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
		320

-13-

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 - 615 620

&lt;210&gt; 10

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 29 GCG

<400> 10  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Ala Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile

-14-

Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
100							105								
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
115							120					125			
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
130						135					140				160
145						150					155				
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
165						170					175				
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
180						185					190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
195						200					205				
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
210						215					220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225						230					235				240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
245						250					255				
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
260						265					270				
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
275						280					285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290						295					300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305						310					315				320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
325						330					335				
Thr	Thr	Gly	Lys	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro	
340						345					350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
355						360					365				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
370						375					380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385						390					395				400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
405						410					415				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420						425					430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435						440					445				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450						455					460				
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465						470					475				480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	
485						490					495				
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
500						505					510				
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
515						520					525				
Arg	Leu	Ala	Arg	Gly	His	Ser	Ieu								
530						535									

&lt;210&gt; 11

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-15-

<220>

<223> Mutant rep protein: rep78 38 GCG

<400> 11  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Ala Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln

-16-

450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val	475	480
465	470	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala	490	495
485		
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val	505	510
500		
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp	525	
515	520	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu	535	540
530		
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys	550	560
545		
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu	555	575
565	570	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr	585	590
580		
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp	600	605
595		
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln	615	620
610		

&lt;210&gt; 12

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 38 GCG

<400> 12		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp	15	
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu	25	30
20		
Lys Glu Trp Glu Leu Ala Pro Asp Ser Asp Met Asp Leu Asn Leu Ile	45	
35	40	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu	60	
50	55	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val	80	
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu	95	
85	90	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile	110	
100	105	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu	125	
115	120	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly	140	
130	135	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys	160	
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu	175	
165	170	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His	190	
180	185	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn	205	
195	200	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr	220	
210	215	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys	240	
225	230	235

-17-

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 - 535

<210> 13

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 39 GCA

<400> 13  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Ala Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile

-18-

	100	105	110													
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu	
		115				120						125				
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly	
		130				135						140				
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys	
		145				150					155			160		
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu	
		165				170						175				
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His	
		180				185						190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn	
		195				200						205				
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr	
		210				215						220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys	
		225				230					235			240		
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala	
		245				250						255				
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys	
		260				265						270				
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln	
		275				280						285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu	
		290				295						300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala	
		305				310					315			320		
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala	
		325				330						335				
Thr	Thr	Gly	Lys	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro		
		340				345						350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp	
		355				360						365				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala	
		370				375						380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg	
		385				390					395			400		
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val	
		405				410						415				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser	
		420				425						430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe	
		435				440						445				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln	
		450				455						460				
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val	
		465				470					475			480		
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala	
		485				490						495				
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val	
		500				505						510				
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp	
		515				520						525				
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu	
		530				535						540				
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys	
		545				550						555			560	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu	
		565				570						575				
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr	
		580				585						590				

-19-

Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595			600			605		
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
							610		615			620			
<210> 14															
<211> 536															
<212> PRT															
<213> Artificial Sequence															
<220>															
<223> Mutant rep protein: rep68 39 GCA															
<400> 14															
Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1						5			10			15			
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
						20		25			30				
Lys	Glu	Trp	Glu	Leu	Pro	Ala	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
						35		40			45				
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
						50		55			60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Ile	Phe	Phe	Val
						65		70		75		80			
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
						85		90			95				
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
						100		105		110					
Arg	Glu	Lys	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu	
						115		120		125					
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	
						130		135		140					
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
						145		150		155		160			
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
						165		170		175					
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
						180		185		190					
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
						195		200		205					
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
						210		215		220					
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
						225		230		235		240			
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
						245		250		255					
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						260		265		270					
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
						275		280		285					
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
						290		295		300					
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
						305		310		315		320			
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						325		330		335					
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
						340		345		350					
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
						355		360		365					
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala

-20-

370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly	Gly Ala Lys Lys Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 15

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 53 GCT

<400> 15		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Ala Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
		240

-21-

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 - 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 16  
<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 53 GCT

-22-

Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
20							25						30		
							40						45		
35															
Glu	Gln	Ala	Pro	Ala	Thr	Val	Ala	Glu	Leu	Gln	Arg	Asp	Phe	Leu	
								55			60				
50															
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Val	
								70			75		80		
65															
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
								85			90		95		
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
								100			105		110		
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
								115			120		125		
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
								130			135		140		
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
								145			150		155		160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
								165			170		175		
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
								180			185		190		
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
								195			200		205		
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210			215		220		
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
								225			230		235		240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								245			250		255		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
								260			265		270		
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
								275			280		285		
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
								290			295		300		
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
								305			310		315		320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
								325			330		335		
Thr	Thr	Gly	lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
								340			345		350		
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
								355			360		365		
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
								370			375		380		
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
								385			390		395		400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
								405			410		415		
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
								420			425		430		
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
								435			440		445		
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
								450			455		460		
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
								465			470		475		480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
								485			490		495		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
								500			505		510		

-23-

```

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
      515          520          525
Arg Leu Ala Arg Gly His Ser Leu
      530          535

<210> 17
<211> 621
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep78 59 GCG

<400> 17
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
    1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
    20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
    35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Ala Gln Arg Asp Phe Leu
    50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
    65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
    85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
   100          105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
   115          120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
   130          135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
   145          150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
   165          170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
   180          185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
   195          200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
   210          215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
   225          230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
   245          250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
   260          265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
   275          280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
   290          295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
   305          310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
   325          330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
   340          345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
   355          360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala

```

-24-

370	375	380
Lys	Val Val Glu Ser Ala Lys Ala Ile Leu	Gly Gly Ser Lys Val Arg
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala	Gln Ile Asp Pro Thr Pro Val	400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys	Ala Val Ile Asp Gly Asn Ser	420
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu	Gln Asp Arg Met Phe Lys Phe	435
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe	Gly Lys Val Thr Lys Gln	450
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala	Lys Asp His Val Val Glu Val	465
465	470	475
Glu His Glu Phe Tyr Val Lys Gly Gly	Ala Lys Lys Arg Pro Ala	480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro	Lys Arg Val Arg Glu Ser Val	500
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu	Ala Ser Ile Asn Tyr Ala Asp	515
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His	Val Gly Met Asn Leu Met Leu	530
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met	Asn Gln Asn Ser Asn Ile Cys	545
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu	Glu Cys Phe Pro Val Ser Glu	560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys	Ala Tyr Gln Lys Leu Cys Tyr	580
580	585	590
Ile His His Ile Met Gly Lys Val Pro	Asp Ala Cys Thr Ala Cys Asp	595
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys	Ile Phe Glu Gln	610
610	615	620

<210> 18

<211> 536

<212> PRT

<213> Artificial Sequence

<220> -

<223> Mutant rep protein: rep 68 59 GCG

<400> 18

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Ala Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160

-25-

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

<210> 19

<211> 621

<212> PRT

<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 64 GCT

<400> 19  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu

-26-

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 20 25 30  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Ala  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510

-27-

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 20  
 <211> 536  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep68 64 GCT

<400> 20  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Ala  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu

-28-

290	295	300
Asn Gly Tyr Asp Pro Gln	Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala	320
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala	335	335
325	330	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro	350	350
340	345	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp	365	365
355	360	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala	380	380
370	375	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg	400	400
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val	410	415
405		
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser	420	430
425		
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe	440	445
435		
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln	450	460
455		
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val	470	475
465		
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala	485	495
490		
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val	500	510
505		
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp	520	525
515		
Arg Leu Ala Arg Gly His Ser Leu	530	535

<210> 21

<211> 621

<212> PRT

<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 74 GCG

<400> 21		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp	15	15
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu	30	30
20	25	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile	45	45
35	40	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu	60	60
50	55	
Thr Glu Trp Arg Arg Val Ser Lys Ala Ala Glu Ala Leu Phe Phe Val	80	80
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu	95	95
85	90	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile	110	110
100	105	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu	125	125
115	120	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly	140	140
130	135	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys	155	160
145	150	

-29-

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
                   165                  170                  175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
                   180                  185                  190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
                   195                  200                  205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
                   210                  215                  220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
                   225                  230                  235                  240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
                   245                  250                  255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
                   260                  265                  270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
                   275                  280                  285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
                   290                  295                  300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
                   305                  310                  315                  320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
                   325                  330                  335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
                   340                  345                  350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
                   355                  360                  365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
                   370                  375                  380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
                   385                  390                  395                  400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
                   405                  410                  415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
                   420                  425                  430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
                   435                  440                  445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
                   450                  455                  460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
                   465                  470                  475                  480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
                   485                  490                  495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
                   500                  505                  510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
                   515                  520                  525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
                   530                  535                  540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
                   545                  550                  555                  560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
                   565                  570                  575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
                   580                  585                  590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
                   595                  600                  605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
                   610                  615                  620

-30-

<212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep68 74 GCG

```

<400> 22
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
   1          5          10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
   20         25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
   35         40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
   50         55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Ala Glu Ala Leu Phe Phe Val
   65         70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
   85         90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
  100        105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
  115        120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
  130        135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
  145        150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
  165        170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
  180        185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
  195        200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
  210        215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
  225        230         235         240
Gln Trp Ile Gin Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
  245        250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
  260        265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
  275        280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
  290        295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
  305        310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
  325        330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
  340        345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
  355        360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
  370        375         380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
  385        390         395         400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
  405        410         415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
  420        425         430

```

-31-

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

<210> 23

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 86 GCG

<400> 23  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Ala Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu

-32-

290	295	300													
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305						310				315					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						325				330					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
						340				345					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
						355				360					365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
						370				375					380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
						385				390					400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
						405				410					415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
						420				425					430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
						435				440					445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
						450				455					460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
						465				470					480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
						485				490					495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
						500				505					510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
						515				520					525
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
						530				535					540
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
						545				550					560
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
						565				570					575
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
						580				585					590
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
						595				600					605
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
						610				615					620

&lt;210&gt; 24

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 86 GCG

&lt;400&gt; 24

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1						5				10					15
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
						20				25					30
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
						35				40					45
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
						50				55					60
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
						65				70					80

-33-

<210> 25  
<211> 621

-34-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 88 GCC

<400> 25  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Ala Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
420 425 430

-35-

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

&lt;210&gt; 26

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein:rep 68 88 GCC

&lt;400&gt; 26

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Ala Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr

-36-

210	215	220
Met Glu Leu Val Gly Trp	Leu Val Asp Lys Gly Ile Thr Ser Glu Lys	240
230	235	
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		255
245	250	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		270
260	265	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		285
275	280	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		300
290	295	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		320
305	310	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		335
325	330	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		350
340	345	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		365
355	360	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		380
370	375	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		400
385	390	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		415
405	410	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		430
420	425	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		445
435	440	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		460
450	455	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		480
465	470	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		495
485	490	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		510
500	505	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		525
515	520	
Arg Leu Ala Arg Gly His Ser Leu		535
530		

<210> 27

<211> 621

<212> PRT

<213> Artificial Sequence

<220> Mutant rep protein: rep 78 101 GCA

<400> 27		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		15
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		30
20	25	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		45
35	40	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		60
50	55	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		80
65	70	

-37-

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
       85                 90                 95  
 Thr Thr Gly Val Ala Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
       100                105                110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
       115                120                125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
       130                135                140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
       145                150                155                160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
       165                170                175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
       180                185                190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
       195                200                205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
       210                215                220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
       225                230                235                240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
       245                250                255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
       260                265                270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
       275                280                285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
       290                295                300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
       305                310                315                320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
       325                330                335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
       340                345                350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
       355                360                365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
       370                375                380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Ser Lys Val Arg  
       385                390                395                400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
       405                410                415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
       420                425                430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
       435                440                445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
       450                455                460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
       465                470                475                480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
       485                490                495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
       500                505                510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
       515                520                525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
       530                535                540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
       545                550                555                560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu

-38-

Ser Gln Pro Val Ser Val Val Lys	565	Ala Tyr Gln Lys	570	Leu Cys Tyr	575
580	585	590			
Ile His His Ile Met Gly Lys Val	595	Pro Asp Ala Cys Thr	600	Ala Cys Asp	605
600	605				
Leu Val Asn Val Asp Leu Asp Asp Cys	610	Ile Phe Glu Gln	615	620	
<210> 28					
<211> 536					
<212> PRT					
<213> Artificial Sequence					
<220>					
<223> Mutant rep protein: rep68 101 GCA					
<400> 28					
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp					
1	5	10	15		
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu					
20	25	30			
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile					
35	40	45			
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu					
50	55	60			
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val					
65	70	75	80		
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu					
85	90	95			
Thr Thr Gly Val Ala Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile					
100	105	110			
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu					
115	120	125			
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly					
130	135	140			
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys					
145	150	155	160		
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu					
165	170	175			
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His					
180	185	190			
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn					
195	200	205			
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr					
210	215	220			
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys					
225	230	235	240		
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala					
245	250	255			
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys					
260	265	270			
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln					
275	280	285			
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu					
290	295	300			
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala					
305	310	315	320		
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala					
325	330	335			
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro					
340	345	350			

-39-

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
           355                 360                 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
           370                 375                 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
           385                 390                 395                 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
           405                 410                 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
           420                 425                 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
           435                 440                 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
           450                 455                 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
           465                 470                 475                 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
           485                 490                 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
           500                 505                 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
           515                 520                 525  
 Arg Leu Ala Arg Gly His Ser Leu  
           530

<210> 29

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 124 GCC

<400> 29

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                 5                 10                 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                 25                 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                 40                 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                 55                 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                 70                 75                 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                 90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                 105                 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ala Glu Pro Thr Leu  
     115                 120                 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
     130                 135                 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145                 150                 155                 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
     165                 170                 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180                 185                 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195                 200                 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr

-40-

210	215	220
Met Glu Leu Val Gly Trp	Leu Val Asp Lys	Gly Ile Thr Ser Glu Lys
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr	Ile Ser Phe Asn Ala Ala	240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys	Ala Ala Leu Asp Asn Ala Gly Lys	260
265	270	275
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr	Leu Val Gly Gln Gln	280
285	290	295
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr	Lys Ile Leu Glu Leu	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val	Phe Leu Gly Trp Ala	310
305	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp	Leu Phe Gly Pro Ala	325
330	335	340
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile	Ala His Thr Val Pro	345
350	355	360
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe	Pro Phe Asn Asp	365
370	375	380
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu	Gly Lys Met Thr Ala	385
390	395	400
Lys Val Val Glu Ser Ala Lys Ala Ile Leu	Gly Gly Ser Lys Val Arg	405
410	415	420
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile	Asp Pro Thr Pro Val	425
430	435	440
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile	Asp Gly Asn Ser	445
450	455	460
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	465
470	475	480
Glu Leu Thr Arg Arg Ile Asp His Asp Phe Gly	Lys Val Thr Lys Gln	485
490	495	500
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val	505
510	515	520
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile	Asn Tyr Ala Asp	525
530	535	540
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly	Met Asn Leu Met Leu	545
550	555	560
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn	Ser Asn Ile Cys	565
570	575	580
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe	Pro Val Ser Glu	585
590	595	600
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln	Lys Leu Cys Tyr	605
610	615	620

&lt;210&gt; 30

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 124 GCC

&lt;400&gt; 30

-41-

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1					5				10					15	
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
					20				25					30	
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
					35			40				45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
					50			55			60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Val	
					65			70			75			80	
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
					85			90				95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
					100			105				110			
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ala	Glu	Pro	Thr	Leu
					115			120			125				
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
					130			135			140				
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
					145			150			155			160	
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
					165			170				175			
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
					180			185			190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
					195			200			205				
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
					210			215			220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
					225			230			235			240	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Asn	Ala	Ala	
					245			250			255				
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					260			265			270				
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					275			280			285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
					290			295			300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
					305			310			315			320	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					325			330			335				
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					340			345			350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					355			360			365				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370			375			380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					385			390			395			400	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					405			410			415				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					420			425			430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					435			440			445				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
					450			455			460				
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					465			470			475			480	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	

-42-

485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 31

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 125 GCG

<400> 31

1	10	15	
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Ala Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	

-43-

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     355                       360                       365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     370                       375                       380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     385                       390                       395                       400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     405                       410                       415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     420                       425                       430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435                       440                       445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450                       455                       460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465                       470                       475                       480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485                       490                       495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     500                       505                       510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     515                       520                       525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
     530                       535                       540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
     545                       550                       555                       560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
     565                       570                       575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
     580                       585                       590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
     595                       600                       605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
     610                       615                       620

&lt;210&gt; 32

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 125 GCG

<400> 32  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1               5                       10                       15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                       25                       30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                       40                       45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                       55                       60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                       70                       75                       80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                       90                       95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                       105                       110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Ala Pro Thr Leu  
     115                       120                       125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly

-44-

130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala.		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Ileu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 33

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 127 GCT

&lt;400&gt; 33

-45-

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Ala Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala

-46-

485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 34

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 127 GCT

&lt;400&gt; 34

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Ala Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270

-47-

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 35

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 132 GCC

&lt;400&gt; 35

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Ala Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly

-48-

130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

-49-

```

<210> 36
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 132 GCC

<400> 36
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
 1           5          10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
 20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
 35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
 50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
 85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Ala Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355         360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370         375         380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385         390         395         400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405         410         415

```

-50-

Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420							425						430		
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435							440						445		
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450							455						460		
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							470						475		
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
													495		
485													490		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
													505		
500													510		
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
													515		
													520		
515													525		
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
													530		
													535		

&lt;210&gt; 37

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 140 GCC

&lt;400&gt; 37

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1						5			10				15		
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
									20				25		30
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
									35				40		45
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
									50				55		60
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65									70				75		80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
									85				90		95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
									100				105		110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
								115				120		125	
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Ala	Ala	Gly	Gly	
								130				135		140	
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145								150				155		160	
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
								165				170		175	
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
								180				185		190	
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
								195				200		205	
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210				215		220	
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225								230				235		240	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								245				250		255	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
								260				265		270	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln

-51-

275	280	285
Pro Val Glu Asp Ile Ser Ser	Asn Arg Ile Tyr	Lys Ile Leu Glu Leu
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val	Phe Leu Gly Trp Ala	
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		480
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 38

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 140 GCC

&lt;400&gt; 38

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		15
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60

-52-

Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Ala Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

-53-

<210> 39  
<211> 621  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Mutant rep protein: rep78 161 GCC  
  
<400> 39  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Ala Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415

-54-

Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420							425						430		
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435							440						445		
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450							455						460		
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							470						475		
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485						490		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500						505		
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515						520		
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
							530						535		
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
							545						550		
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
							565						570		
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
							580						585		
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595						600		
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
							610						615		
													620		

&lt;210&gt; 40

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 161 GCC

&lt;400&gt; 40

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1						5			10			15			
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
						20			25			30			
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
						35			40			45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
						50			55			60			
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
						65			70			75			80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
						85			90			95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
						100			105			110			
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
						115			120			125			
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	
						130			135			140			
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
						145			150			155			160
Ala	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
						165			170			175			
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
						180			185			190			
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn

-55-

195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
400		
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
480		
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 41

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 163 GCT

<400> 41

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		15
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60

-56-

Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70					75					80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
										90					95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
										105					110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
										120					125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
										135					140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145										150					160
Thr	Gln	Ala	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
										165					175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
										180					190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
										195					205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
										210					220
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225										230					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
										245					255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
										260					270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
										275					285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
										290					300
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305										310					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
										325					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
										340					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
										355					365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
										370					380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385										390					400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
										405					415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
										420					430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
										435					445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
										450					460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465										470					480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
										485					495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
										500					510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
										515					525
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
										530					540
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys

-57-

545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu	Cys Phe Pro Val Ser Glu		
565	570		575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

&lt;210&gt; 42

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 163 GCT

<400> 42			
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Ala Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	

-58-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
                   340                 345                 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
                   355                 360                 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
                   370                 375                 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
                   385                 390                 395                 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
                   405                 410                 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
                   420                 425                 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
                   435                 440                 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
                   450                 455                 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
                   465                 470                 475                 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
                   485                 490                 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
                   500                 505                 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
                   515                 520                 525  
 Arg Leu Ala Arg Gly His Ser Leu  
                   530                 535

&lt;210&gt; 43

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 175 GCT

<400> 43  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1              5                 10                 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                 25                 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                 40                 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                 55                 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                 70                 75                 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                 90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                 105                 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115                 120                 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130                 135                 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145                 150                 155                 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Ala Leu  
     165                 170                 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180                 185                 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn

-59-

195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 44

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 175 GCT

-60-

<400> 44  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
   1               5               10               15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
   20              25              30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
   35              40              45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
   50              55              60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
   65              70              75              80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
   85              90              95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
   100             105             110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
   115             120             125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
   130             135             140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
   145             150             155             160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Ala Leu  
   165             170             175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
   180             185             190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
   195             200             205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
   210             215             220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
   225             230             235             240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
   245             250             255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
   260             265             270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
   275             280             285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
   290             295             300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
   305             310             315             320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
   325             330             335  
 Thr Thr Gly Lys Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
   340             345             350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
   355             360             365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
   370             375             380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
   385             390             395             400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
   405             410             415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
   420             425             430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
   435             440             445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
   450             455             460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
   465             470             475             480

-61-

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

<210> 45  
 <211> 621  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep78 193 GCG

<400> 45  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Ala Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro

-62-

Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
340							345						350		
355							360						365		
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
370							375						380		
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385							390						395		
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
											405			415	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
											420			430	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
											435			445	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
											450			460	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
											465			475	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
											485			490	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
											500			505	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
											515			520	
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
											530			535	
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
											545			550	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
											565			570	
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
											580			585	
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
											595			600	
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
											610			615	
														620	

&lt;210&gt; 46

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 193 GCG

&lt;400&gt; 46

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1						5			10				15		
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
									20			25		30	
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
									35			40		45	
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
									50			55		60	
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
									65			70		75	
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
									85			90		95	
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
									100			105		110	
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
									115			120		125	

-63-

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Ala Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 47

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 196 GCC

-64-

<400> 47  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Ala Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480

-65-

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

&lt;210&gt; 48

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 196 GCC

<400> 48  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Ala Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys

-66-

	260	265	270
Ile Met Ser	Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		

<210> 49

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 197 GCC

<400> 49

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	

-67-

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ala Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln

-68-

610

615

620

```

<210> 50
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 197 GCC

<400> 50
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
      5          10          15
      1           5           10           15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
      20          25          30
      20           25           30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
      35          40          45
      35           40           45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
      50          55          60
      50           55           60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
      65          70          75          80
      65           70           75           80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
      85          90          95
      85           90           95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
      100         105         110
      100          105          110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
      115         120         125
      115          120          125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
      130         135         140
      130          135          140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
      145         150         155         160
      145          150          155          160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
      165         170         175
      165          170          175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
      180         185         190
      180          185          190
Leu Thr His Val Ala Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
      195         200         205
      195          200          205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
      210         215         220
      210          215          220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
      225         230         235         240
      225          230          235          240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
      245         250         255
      245          250          255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
      260         265         270
      260          265          270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
      275         280         285
      275          280          285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
      290         295         300
      290          295          300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
      305         310         315         320
      305          310          315          320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
      325         330         335
      325          330          335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
      340         345         350
      340          345          350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
      355         360         365
      355          360          365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
      370         375         380
      370          375          380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
      385         390         395         400
      385          390          395          400

```

-69-

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
                   405                  410                  415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
                   420                  425                  430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
                   435                  440                  445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
                   450                  455                  460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
                   465                  470                  475                  480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
                   485                  490                  495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
                   500                  505                  510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
                   515                  520                  525  
 Arg Leu Ala Arg Gly His Ser Leu  
                   530                  535

&lt;210&gt; 51

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 221 GCA

<400> 51  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                  5                  10                  15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                  25                  30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                  40                  45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                  55                  60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                  70                  75                  80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                  90                  95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                105                  110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115                120                  125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130                135                  140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145                150                  155                  160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
     165                170                  175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180                185                  190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195                200                  205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ala Ala Arg Tyr  
     210                215                  220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     225                230                  235                  240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     245                250                  255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys

-70-

Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
260							265						270		
275							280						285		
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290							295						300		
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310						315		
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325						330		
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
							340						345		
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355						360		
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370						375		
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
							385						390		
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405						410		
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420						425		
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435						440		
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450						455		
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
							465						470		
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485						490		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500						505		
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515						520		
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
							530						535		
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
							545						550		
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
							565						570		
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
							580						585		
Ile	His	Ile	His	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595						600		
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
							610						615		
													620		
<210>	52														
<211>	536														
<212>	PRT														
<213>	Artificial Sequence														
<220>															
<223>	Mutant rep protein:	rep68	221	GCA											
<400>	52														
Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
							1						5		10
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
							20						25		30
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Ile	Ile
							35						40		45

-71-

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ala Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu

-72-

530

535

```

<210> 53
<211> 621
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep78 228 GCG

<400> 53
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
      1          5          10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
      20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
      35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
      50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
      65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
      85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
      100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
      115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
      130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
      145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
      165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
      180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
      195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
      210         215         220
Met Glu Leu Ala Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
      225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
      245         250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
      260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
      275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
      290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
      305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
      325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
      340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
      355         360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
      370         375         380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
      385         390         395         400

```

-73-

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
           405                 410                 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
           420                 425                 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
           435                 440                 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
           450                 455                 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
           465                 470                 475                 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
           485                 490                 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
           500                 505                 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
           515                 520                 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
           530                 535                 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
           545                 550                 555                 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
           565                 570                 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
           580                 585                 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
           595                 600                 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
           610                 615                 620

&lt;210&gt; 54

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 228 GCG

<400> 54  
 Met Glu Leu Ala Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1              5                 10                 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20                 25                 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35              40                 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50              55                 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65              70                 75                 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85              90                 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100             105                 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115             120                 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130             135                 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145             150                 155                 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165             170                 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val

-74-

180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
305	310	315
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
325	330	335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
340	345	350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
355	360	365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
370	375	380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
385	390	395

&lt;210&gt; 55

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 228 GCG

&lt;400&gt; 55

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	

-75-

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Ala Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 56

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 228 GCG

&lt;400&gt; 56

Met Glu Leu Ala Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln

-76-

50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		80
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		160
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Leu Ala Arg Gly His Ser Leu		
305	310	

&lt;210&gt; 57

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 231 GCC

&lt;400&gt; 57

1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	

-77-

Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145				150					155					160	
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
									170					175	
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
									185					190	
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
									200					205	
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
							210	215		220					
Met	Glu	Leu	Val	Gly	Trp	Ala	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225							230		235					240	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							245		250					255	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
							260		265					270	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
							275		280					285	
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
							290		295					300	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310		315					320	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325		330					335	
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
							340		345					350	
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355		360					365	
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370		375					380	
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385							390		395					400	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405		410					415	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420		425					430	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435		440					445	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450		455					460	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							470		475					480	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485		490					495	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500		505					510	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515		520					525	
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
							530		535					540	
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
545							550		555					560	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
							565		570					575	
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
							580		585					590	
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595		600					605	
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
							610		615					620	

-78-

<210> 58  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: 52 231 GCC

<400> 58  
Met Glu Leu Val Gly Trp Ala Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
305 310 315 320  
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
325 330 335  
Phe Thr His Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
340 345 350  
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
355 360 365  
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
385 390 395

<210> 59

-79-

<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 231 GCC

<400> 59  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Ala Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser

-80-

420	425	430
Thr Thr Phe Glu His Gln Gln Pro	Leu Gln Asp Arg Met	Phe Lys Phe
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val	455	460
450		
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val	470	475
465		480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala	485	495
500	490	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	515	Tyr Ala Asp
520	525	
Arg Leu Ala Arg Gly His Ser Leu	530	
	535	

<210> 60

<211> 312

<212> PRT

<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 231 GCC

<400> 60		
Met Glu Leu Val Gly Trp Ala Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala	20	25
20		30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys	35	40
35		45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln	50	55
50		60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu	65	70
65		75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala	85	90
85		95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala	100	105
100		110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro	115	120
115		125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp	130	135
130		140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala	145	150
145		155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg	165	170
165		175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val	180	185
180		190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser	195	200
195		205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe	210	215
210		220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln	225	230
225		235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val	245	250
245		255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala	260	265
260		270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val	275	280
		285

-81-

```

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
290          295          300
Arg Leu Ala Arg Gly His Ser Leu
305          310

<210> 61
<211> 621
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep78 234 GCG

<400> 61
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Ala Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355         360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala

```

-82-

370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 62

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 234 GCG

&lt;400&gt; 62

Met Glu Leu Val Gly Trp Leu Val Asp Ala Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		15
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		80
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
		160

-83-

Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					165				170						175
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					180				185						190
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					195			200			205				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					210			215			220				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
					225			230		235					240
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					245				250						255
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
					260			265							270
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
					275			280			285				
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
					290			295			300				
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
					305			310		315					320
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
					325				330						335
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
					340			345							350
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
					355			360			365				
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
					370			375			380				
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
					385			390			395				

```
<210> 63
<211> 536
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Mutant rep protein: rep68 234 GCG

<400> 63  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu

-84-

	165		170		175
Ser Ala Cys Leu Asn Leu Thr Glu Arg		Lys Arg Leu Val Ala Gln His			
180		185		190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys					
195		200		205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys		Thr Ser Ala Arg Tyr			
210		215		220	
Met Glu Leu Val Gly Trp Leu Val Asp Ala Gly		Ile Thr Ser Glu Lys			
225		230		235	
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala					
245		250		255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys					
260		265		270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln					
275		280		285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu					
290		295		300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala					
305		310		315	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala					
325		330		335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro					
340		345		350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp					
355		360		365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala					
370		375		380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg					
385		390		395	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val					
405		410		415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser					
420		425		430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe					
435		440		445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln					
450		455		460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val					
465		470		475	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala					
485		490		495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val					
500		505		510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp					
515		520		525	
Arg Leu Ala Arg Gly His Ser Leu					
530		535			

<210> 64

<211> 312

<212> PRT

**<213> Artificial Sequence**

<220>

<223> Mutant rep protein: rep40 234 GCG

<400> 64  
Mat. Clu

Met Glu Leu Val Gly Ile Ser Val Asp Ala Gly Ile Ile Ile Ser Val  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30

-85-

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                        40                        45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                        55                        60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                        70                        75                        80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                        90                        95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                       105                       110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115                       120                       125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130                       135                       140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145                       150                       155                       160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                       170                       175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                       185                       190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195                       200                       205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210                       215                       220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225                       230                       235                       240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245                       250                       255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260                       265                       270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275                       280                       285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290                       295                       300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305                       310

&lt;210&gt; 65

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 237 GCC

&lt;400&gt; 65

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                        5                        10                        15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                       25                       30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                       40                       45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                       55                       60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                       70                       75                       80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                       90                       95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                       105                       110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu

-86-

115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Ala Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605

-87-

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610                   615                   620

<210> 66

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 237 GCC

<400> 66

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Ala	Ser	Glu	Lys
1				5					10				15		
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								20		25			30		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						35				40		45			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					50			55		60					
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Glu	Leu	
65					70				75		80				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
					85				90		95				
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					100			105			110				
Thr	Thr	Gly	lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					115			120			125				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					130			135		140					
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
145					150				155		160				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					165				170		175				
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					180				185		190				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					195			200			205				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					210			215		220					
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
225					230				235		240				
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					245				250		255				
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	
					260			265		270					
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
					275			280		285					
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
					290			295		300					
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
305					310				315		320				
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
					325				330		335				
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
					340			345		350					
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
					355			360		365					
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
					370			375		380					
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			

-88-

385	390	395
<210> 67		
<211> 536		
<212> PRT		
<213> Artificial Sequence		
<220>		
<223> Mutant rep protein: rep68 237 GCC		
<400> 67		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1 5 10 15		
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		
20 25 30		
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35 40 45		
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50 55 60		
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65 70 75 80		
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85 90 95		
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100 105 110		
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115 120 125		
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130 135 140		
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145 150 155 160		
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165 170 175		
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180 185 190		
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195 200 205		
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210 215 220		
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Ala Ser Glu Lys		
225 230 235 240		
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245 250 255		
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260 265 270		
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275 280 285		
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290 295 300		
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305 310 315 320		
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325 330 335		
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340 345 350		
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355 360 365		
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370 375 380		
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385 390 395 400		

-89-

```

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
        405          410          415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
        420          425          430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
        435          440          445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
        450          455          460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
        465          470          475          480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
        485          490          495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
        500          505          510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
        515          520          525
Arg Leu Ala Arg Gly His Ser Leu
        530          535

```

&lt;210&gt; 68

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;
&lt;223&gt; Mutant rep protein: rep40 237 GCC

```

<400> 68
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Ala Ser Glu Lys
    1          5          10          15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
    20          25          30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
    35          40          45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
    50          55          60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
    65          70          75          80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
    85          90          95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
    100         105         110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
    115         120         125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
    130         135         140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
    145         150         155         160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
    165         170         175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
    180         185         190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
    195         200         205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
    210         215         220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
    225         230         235         240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
    245         250         255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Arg Pro Ala

```

-90-

-91-

```

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355          360          365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370          375          380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385          390          395          400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405          410          415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
420          425          430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435          440          445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
450          455          460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
465          470          475          480
Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala
485          490          495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500          505          510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515          520          525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530          535          540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545          550          555          560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565          570          575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580          585          590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595          600          605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610          615          620

```

&lt;210&gt; 70

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 250 GCC

```

<400> 70
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
1          5          10          15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Ala Ile Ser Phe Asn Ala Ala
20         25         30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
35         40         45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
50         55         60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
65         70         75         80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
85         90         95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
100        105        110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
115        120        125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp

```

-92-

130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		160
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
305	310	315
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		320
325	330	335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
340	345	350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
355	360	365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
370	375	380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
385	390	395

&lt;210&gt; 71

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 250 GCC

<400> 71		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		15
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		80
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140

-93-

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Ala Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

<210> 72

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 250 GCC

<400> 72

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys

-94-

1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Ala Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Leu Ala Arg Gly His Ser Leu			
305	310		

&lt;210&gt; 73

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 258 GCC

&lt;400&gt; 73

1	5	10	15
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
20	25	30	
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
35	40	45	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
50	55	60	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
65	70	75	80
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
85	90	95	
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			

-95-

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Ala Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr

-96-

580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

<210> 74  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep52 258 GCC

<400> 74	<400> 74	<400> 74	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Ala Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	345	350	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
355	360	365	

-97-

```

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
 370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385          390          395

<210> 75
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 258 GCC

<400> 75
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
 1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
 20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
 35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
 50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
 85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Ala Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355         360         365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala

```

-98-

370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	395 400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile	Asp Pro Thr Pro Val	
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val	Ile Asp Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp	Arg Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe	Gly Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys	Asp His Val Val Glu Val	
465	470	475 480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly	Ala Lys Lys Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg	Val Arg Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser	Ile Asn Tyr Ala Asp	
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 76

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 258 GCC

&lt;400&gt; 76

Met Glu Leu Val Gly Trp Leu Val Asp Lys	Gly Ile Thr Ser Glu Lys
1	5 10 15
Gln Trp Ile Gln Glu Asp Gln Ala Ser	Tyr Ile Ser Phe Asn Ala Ala
20	25 30
Ser Ala Ser Arg Ser Gln Ile Lys Ala Ala	Leu Asp Asn Ala Gly Lys
35	40 45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp	Tyr Leu Val Gly Gln Gln
50	55 60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile	Tyr Lys Ile Leu Glu Leu
65	70 75 80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser	Val Phe Leu Gly Trp Ala
85	90 95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile	Trp Leu Phe Gly Pro Ala
100	105 110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala	Ile His Thr Val Pro
115	120 125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu	Asn Phe Pro Phe Asn Asp
130	135 140
Cys Val Asp Lys Met Val Ile Trp Trp Glu	Glu Gly Lys Met Thr Ala
145	150 155 160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu	Gly Ser Lys Val Arg
165	170 175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln	Ile Asp Pro Thr Pro Val
180	185 190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala	Val Ile Asp Gly Asn Ser
195	200 205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln	Asp Arg Met Phe Lys Phe
210	215 220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe	Gly Lys Val Thr Lys Gln
225	230 235 240

-99-

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

<210> 77  
 <211> 621  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: 78 260 GCG

<400> 77  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Ala Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala

-100-

325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His	Thr Val Pro	
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro	Phe Asn Asp	
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys	Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser	Lys Val Arg	
385	390	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro	Thr Pro Val	
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met	Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val	Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val	Val Glu Val	
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys	Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn	Leu Met Leu	
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser	Asn Ile Cys	
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro	Val Ser Glu	
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys	Leu Cys Tyr	
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr	Ala Cys Asp	
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 78

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 260 GCG

&lt;400&gt; 78

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser	Glu Lys		
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe	Asn Ala Ala		
20	25	30	
Ser Asn Ser Ala Ser Gln Ile Lys Ala Ala Leu Asp Asn	Ala Gly Lys		
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val	Gly Gln Gln		
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile	Leu Glu Leu		
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu	Gly Trp Ala		
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe	Gly Pro Ala		
100	105	110	

-101-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

&lt;210&gt; 79

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 260 GCG

&lt;400&gt; 79

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu

-102-

115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Ala Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 80

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 260 GCG

-103-

<400> 80  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1              5                 10                 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20                 25                 30  
 Ser Asn Ser Ala Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                 40                 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                 55                 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                 70                 75                 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                 90                 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                105                110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115                120                125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130                135                140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145                150                155                160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                170                175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                185                190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195                200                205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210                215                220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225                230                235                240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245                250                255  
 Glu His Glu Phe Tyr Val Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260                265                270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275                280                285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290                295                300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305                310

&lt;210&gt; 81

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 263 GCC

<400> 81  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1              5                 10                 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                 25                 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                 40                 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                 55                 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                 70                 75                 80

-104-

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
       85                 90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
       100                105                110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
       115                120                125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
       130                135                140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
       145                150                155                160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
       165                170                175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
       180                185                190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
       195                200                205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
       210                215                220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
       225                230                235                240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
       245                250                255  
 Ser Asn Ser Arg Ser Gln Ala Lys Ala Ala Leu Asp Asn Ala Gly Lys  
       260                265                270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
       275                280                285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
       290                295                300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
       305                310                315                320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
       325                330                335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
       340                345                350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
       355                360                365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
       370                375                380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
       385                390                395                400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
       405                410                415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
       420                425                430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
       435                440                445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
       450                455                460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
       465                470                475                480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
       485                490                495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
       500                505                510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
       515                520                525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
       530                535                540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
       545                550                555                560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu

-105-

Ser Gln Pro Val	565	Ser Val Val Lys	570	Lys Ala Tyr Gln Lys	575
580	585	Leu Cys Tyr	590		
Ile His His Ile Met	595	Gly Lys Val Pro Asp Ala Cys Thr	600	Ala Cys Asp	605
Leu Val Asn Val Asp	610	Leu Asp Asp Cys Ile Phe Glu	615	Gln	620
<210> 82					
<211> 397					
<212> PRT					
<213> Artificial Sequence					
<220>					
<223> Mutant rep protein: rep52 263 GCC					
<400> 82					
Met Glu Leu Val Gly Trp	1	Leu Val Asp Lys	10	Gly Ile Thr Ser Glu Lys	15
Trp	5	15			
Gln Trp Ile Gln Glu Asp	20	Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala	25	30	
Ser Asn Ser Arg Ser Gln Ala	35	Lys Ala Ala Leu Asp Asn Ala Gly Lys	40	45	
Ile Met Ser Leu Thr Lys	50	Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg	65	Ile Tyr Lys Ile Leu Glu Leu	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala	85	Ser Val Phe Leu Gly Trp Ala	90	95	
Thr Lys Phe Gly Lys Arg Asn Thr	100	Ile Trp Leu Phe Gly Pro Ala	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala	115	His Thr Val Pro	120	125	
Phe Tyr Gly Cys Val Asn Trp	130	Thr Asn Glu Asn Phe Pro Phe Asn Asp	135	140	
Cys Val Asp Lys Met Val	145	Ile Trp Trp Glu Gly Lys Met Thr Ala	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	165	Gly Ser Lys Val Arg	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp	180	Pro Thr Pro Val	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile	195	Asp Gly Asn Ser	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	210	Met Phe Lys Phe	215	220	
Glu Leu Thr Arg Arg	225	Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln	230	235	240
Glu Val Lys Asp Phe	245	Arg Trp Ala Lys Asp His Val Val Glu Val	250	255	
Glu His Glu Phe Tyr Val Lys	260	Lys Gly Ala Lys Lys Arg Pro Ala	265	270	
Pro Ser Asp Ala Asp Ile Ser	275	Glu Pro Lys Arg Val Arg Glu Ser Val	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile	290	Asn Tyr Ala Asp	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val	305	Gly Met Asn Leu Met Leu	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser	325	Asn Ile Cys	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu	340	Glu Cys Phe Pro Val Ser Glu	345	350	

-106-

```

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
355          360          365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385          390          395

<210> 83
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 263 GCC

<400> 83
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Asn Ser Arg Ser Gln Ala Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp

```

-107-

355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly	Lys Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly	Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro	Thr Pro Val	
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val	
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly	Gly Ala Lys Lys Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 84

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 263 GCC

&lt;400&gt; 84

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20	25	30
Ser Asn Ser Arg Ser Gln Ala Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220

-108-

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

&lt;210&gt; 85

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 264 GCG

<400> 85  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala

-109-

305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
565	570	575	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

<210> 86

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 264 GCG

<400> 86

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	

-110-

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

<210> 87

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 264 GCG

<400> 87

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile

-111-

100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Ala Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 88

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-112-

<220>  
<223> Mutant rep protein: rep40 264 GCG

<400> 88  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1              5                 10                15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20              25                30  
 Ser Asn Ser Arg Ser Gln Ile Ala Ala Ala Leu Asp Asn Ala Gly Lys  
     35              40                45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50              55                60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65              70                75                80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85              90                95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100            105              110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115            120              125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130            135              140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145            150              155                160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165            170              175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180            185              190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195            200              205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210            215              220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225            230              235                240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245            250              255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260            265              270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275            280              285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290            295              300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305            310

<210> 89  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 334 GCG

<400> 89  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1              5                 10                15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20              25                30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35              40                45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu

-113-

50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala	Leu Phe Phe Val	
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His	Val Leu Val Glu	80
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe	Leu Ser Gln Ile	
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile	Glu Pro Thr Leu	
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly	Ala Gly Gly Gly	
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr	Leu Leu Pro Lys	
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met	Glu Gln Tyr Leu	160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu	Val Ala Gln His	
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys	Glu Asn Gln Asn	
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr	Ser Ala Arg Tyr	
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile	Thr Ser Glu Lys	
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser	Phe Asn Ala Ala	240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp	Asn Ala Gly Lys	
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu	Val Gly Gln Gln	
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys	Ile Leu Glu Leu	
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe	Leu Gly Trp Ala	
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu	Phe Ala Pro Ala	320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala	His Thr Val Pro	
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe	Pro Phe Asn Asp	
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly	Lys Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly	Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp	Pro Thr Pro Val	400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile	Asp Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys	Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val	
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys	Lys Arg Pro Ala	480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val	Arg Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile	Asn Tyr Ala Asp	
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met	Asn Leu Met Leu	
530	535	540

-114-

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 90  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep52 334 GCG

<400> 90  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Ala Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys

-115-

Phe	Thr	His	Gly	325	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
				340					345						350	
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr	
									355	360				365		
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp	
								370	375		380					
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln				
								385	390		395					

<210> 91

<211> 536

<212> PRT

<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 334 GCG

91																	
Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp		
1						5			10						15		
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu		
								20	25					30			
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile		
								35	40			45					
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu		
								50	55		60						
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val		
								65	70		75		80				
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu		
								85	90		95						
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile		
								100	105		110						
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu		
								115	120		125						
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly			
							130	135		140							
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys		
							145	150		155		160					
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu		
							165	170		175							
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His		
							180	185		190							
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn		
							195	200		205							
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr		
							210	215		220							
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys		
							225	230		235		240					
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala		
							245	250		255							
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys		
							260	265		270							
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln		
							275	280		285							
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu		
							290	295		300							
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala		
							305	310		315		320					
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Ala	Pro	Ala		
							325	330		335							

-116-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 92

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 334 GCG

<400> 92  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Ala Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser

-117-

195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Leu Ala Arg Gly His Ser Leu		
305	310	

<210> 93  
 <211> 621  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep78 335 GCT

400	93	
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285

-118-

Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290					295						300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305					310					315					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Ala	Ala
						325				330					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
						340				345					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					355				360						365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370				375						380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385						390				395					400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
						405				410					415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
						420				425					430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					435				440						445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
					450				455						460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465						470				475					480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	
						485				490					495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
					500				505						510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
					515				520						525
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
					530				535						540
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
545						550				555					560
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
						565				570					575
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
						580				585					590
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
					595				600						605
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
					610				615						620

&lt;210&gt; 94

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 335 GCT

&lt;400&gt; 94

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1					5				10						15
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
						20			25						30
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						35			40						45
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					50				55						60
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu

-119-

65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Ala Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	345	350	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
355	360	365	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
370	375	380	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
385	390	395	

&lt;210&gt; 95

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: 68 335 GCT

&lt;400&gt; 95

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		*	
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80

-120-

Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
85									90						95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
100								105						110	
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
115							120					125			
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
130							135				140				
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145						150				155					160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
165							170				175				
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
180							185				190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
195							200				205				
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
210							215				220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225						230				235					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
245							250				255				
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
260							265				270				
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
275							280				285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Glu	Leu	
290							295				300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310				315				320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Ala	Ala
325							330				335				
Thr	Thr	Gly	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro	
340							345				350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
355							360				365				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
370							375				380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385							390				395				400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
405							410				415				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420							425				430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435							440				445				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450							455				460				
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							470				475				480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
485							490				495				
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
500							505				510				
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
515							520				525				
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
530							535								

-121-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 335 GCT

<400> 96  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Ala Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

<210> 97

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 337 GCT

<400> 97

Thr Ala G

1  
—

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Ile Val Ala Glu  
20 25 30

-122-

Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
35					40						45				
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
50						55					60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65						70				75			80		
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
									85	90		95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
								100	105		110				
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
							115	120		125					
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
						130	135		140						
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145						150				155			160		
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
							165		170		175				
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
							180		185		190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
							195	200		205					
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
							210	215		220					
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225							230			235			240		
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							245		250		255				
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						260	265		270						
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
							275	280		285					
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Glu	Leu	
							290	295		300					
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310			315			320		
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325		330		335				
Ala	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
							340	345		350					
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355	360		365					
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370	375		380					
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385							390		395		400				
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405		410		415				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420	425		430					
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435	440		445					
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450	455		460					
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							470		475		480				
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	
							485		490		495				
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500	505		510					
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp

-123-

515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 98

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 337 GCT

&lt;400&gt; 98

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Ala Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	

-124-

```

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
305          310          315          320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
325          330          335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
340          345          350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
355          360          365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385          390          395

```

&lt;210&gt; 99

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 68 337 GCT

&lt;400&gt; 99

```

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1          5          10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala

```

-125-

305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Ala Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		

<210> 100

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 337 GCT

<400> 100

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Ala Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	

-126-

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

&lt;210&gt; 101

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 78 341 GCC

&lt;400&gt; 101

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys

-127-

260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
320	325	330
Thr Thr Gly Lys Ala Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620
<210> 102		
<211> 397		
<212> PRT		
<213> Artificial Sequence		
<220>		
<223> Mutant rep protein: rep52 341 GCC		
<400> 102		
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45

-128-

Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
50					55						60				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
65					70					75				80	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
					85					90				95	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					100					105				110	
Thr	Thr	Gly	Lys	Ala	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					115					120				125	
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					130					135				140	
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
145					150					155				160	
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					165					170				175	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					180					185				190	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					195					200				205	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					210					215				220	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
225					230					235				240	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					245					250				255	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
					260					265				270	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
					275					280				285	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
					290					295				300	
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
305					310					315				320	
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
					325					330				335	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
					340					345				350	
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
					355					360				365	
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
					370					375				380	
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
					385					390				395	

<210> 103  
<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 341 GCC

<400> 103  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Leu Gln Arg Asp Phe Leu

-129-

50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		80
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Ala Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

-130-

<210> 104  
<211> 312  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 341 GCC

<400> 104  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Ala Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Leu Ala Arg Gly His Ser Leu  
305 310

<210> 105  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 342 GCC

<400> 105  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15

-131-

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Ala Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val

-132-

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp	500	505	510
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu	530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys	545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu	565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr	580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp	595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln	610	615	620

&lt;210&gt; 106

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 342 GCC

&lt;400&gt; 106

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys	1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala	20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys	35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln	50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu	65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala	85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala	100	105	110	
Thr Thr Gly Lys Thr Ala Ile Ala Glu Ala Ile Ala His Thr Val Pro	115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp	130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala	145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg	165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val	180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser	195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe	210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln	225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val	245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala	260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val	275	280	285	

-133-

```

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
290          295          300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
305          310          315          320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
325          330          335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
340          345          350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
355          360          365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385          390          395

<210> 107
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 342 GCC

<400> 107
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1          5          10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20         25         30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35         40         45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50         55         60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65         70         75         80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85         90         95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100        105        110
Arg Glu Lys Ile Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115        120        125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130        135        140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145        150        155        160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165        170        175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180        185        190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195        200        205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210        215        220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225        230        235        240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245        250        255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260        265        270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275        280        285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu

```

-134-

290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val	Phe Leu Gly Trp Ala	
305	310	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp	Leu Phe Gly Pro Ala	
325	330	335
Thr Thr Gly Lys Thr Ala Ile Ala Glu Ala Ile	Ala His Thr Val Pro	
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn	Phe Pro Phe Asn Asp	
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu	Gly Lys Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile	Asp Pro Thr Pro Val	
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val	Ile Asp Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp	Arg Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp	His Val Val Glu Val	
465	470	480
Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys	Lys Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg	Val Arg Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser	Ile Asn Tyr Ala Asp	
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 108

<211> 312

<212> PRT

### <213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 342 GCG

<400> 108  
Met Glu L

Met	Glu	Leu	Val	Gly	115	Ile	Val	Asp	Lys	Cys	116	Ile	Ala	Ala	15
1				5					10						
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
				20					25					30	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
				35					40				45		
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
				50					55			60			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
				65					70			75			80
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
				85					90					95	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
				100					105					110	
Thr	Thr	Gly	Lys	Thr	Ala	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
				115					120				125		
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
				130					135			140			
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
				145					150			155			160

-135-

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                170                175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                185                190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195                200                205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210                215                220  
 Glu Leu Thr Arg Arg Leu Asp His Phe Gly Lys Val Thr Lys Gln  
     225                230                235                240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245                250                255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
     260                265                270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275                280                285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290                295                300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305                310

&lt;210&gt; 109

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 347 GCA

<400> 109  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                5                10                15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                25                30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                40                45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                55                60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                70                75                80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                90                95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100               105               110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115               120               125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130               135               140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145               150               155                160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
     165               170               175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180               185               190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195               200               205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
     210               215               220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     225               230               235                240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala

-136-

Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
245								250					255		
260							265					270			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
275							280					285			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290							295					300			
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310					315			320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325					330			335
Thr	Thr	Gly	Thr	Asn	Ile	Ala	Glu	Ala	Ala	Ala	His	Thr	Val	Pro	
							340					345			350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355					360			365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370					375			380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
							385					390			395
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405					410			415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420					425			430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435					440			445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450					455			460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
							465					470			475
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485					490			495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500					505			510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	'Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515					520			525
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
							530					535			540
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
							545					550			560
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
							565					570			575
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
							580					585			590
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
							595					600			605
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
							610					615			620

&lt;210&gt; 110

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 347 GCA

&lt;400&gt; 110

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1						5				10			15		
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							20					25			30

-137-

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                        40                        45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                        55                        60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                        70                        75                        80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                        90                        95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                       105                       110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ala Ala His Thr Val Pro  
     115                       120                       125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130                       135                       140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145                       150                       155                       160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                       170                       175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                       185                       190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195                       200                       205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210                       215                       220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225                       230                       235                       240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245                       250                       255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260                       265                       270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275                       280                       285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290                       295                       300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
     305                       310                       315                       320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
     325                       330                       335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
     340                       345                       350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
     355                       360                       365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
     370                       375                       380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
     385                       390                       395

&lt;210&gt; 111

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 347 GCA

&lt;400&gt; 111

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                        5                        10                        15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                       25                       30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile

-138-

Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
35															45
50								55		60					
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65								70		75					80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
								85		90					95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
								100		105					110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
								115		120					125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
								130		135					140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
								145		150					160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
								165		170					175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
								180		185					190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
								195		200					205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210		215					220
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
								225		230					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
								245		250					255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
								260		265					270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
								275		280					285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
								290		295					300
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
								305		310					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
								325		330					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ala	Ala	His	Thr	Val	Pro
								340		345					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
								355		360					365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
								370		375					380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
								385		390					400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
								405		410					415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
								420		425					430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
								435		440					445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
								450		455					460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
								465		470					480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
								485		490					495
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
								500		505					510
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
								515		520					525

-139-

Arg Leu Ala Arg Gly His Ser Leu  
530 535

<210> 112  
<211> 312  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 347 GCA

<400> 112  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ala Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Leu Ala Arg Gly His Ser Leu  
305 310

<210> 113  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep 78 350 AAT

-140-

<400> 113  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Asn Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480

-141-

```

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
485          490          495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500          505          510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515          520          525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530          535          540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545          550          555          560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565          570          575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580          585          590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595          600          605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610          615          620

```

<210> 114  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep52 350 AAT

```

<400> 114
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
1           5           10          15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
20          25          30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
35          40          45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
50          55          60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
65          70          75          80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
85          90          95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
100         105         110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Asn Val Pro
115         120         125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
130         135         140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
145         150         155         160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
165         170         175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
180         185         190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
195         200         205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
210         215         220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
225         230         235         240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
245         250         255
Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala

```

-142-

260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val		
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	345	350	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
355	360	365	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
370	375	380	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
385	390	395	

&lt;210&gt; 115

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 350 AAT

&lt;400&gt; 115

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	

-143-

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Asn Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 116

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 350 AAT

&lt;400&gt; 116

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Asn Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp

-144-

130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Leu Ala Arg Gly His Ser Leu			
305	310		

<210> 117

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 350 GCT

<400> 117

1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	

-145-

```

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225      230      235      240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245      250      255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260      265      270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275      280      285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290      295      300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305      310      315      320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325      330      335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
340      345      350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355      360      365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370      375      380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385      390      395      400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405      410      415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
420      425      430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435      440      445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
450      455      460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
465      470      475      480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
485      490      495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500      505      510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515      520      525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530      535      540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545      550      555      560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565      570      575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580      585      590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595      600      605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610      615      620

```

&lt;210&gt; 118

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 350 GCT

&lt;400&gt; 118

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys

-146-

1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	345	350	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
355	360	365	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
370	375	380	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
385	390	395	

&lt;210&gt; 119

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 350 GCT

&lt;400&gt; 119

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15

-147-

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                       25                       30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                       40                       45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                       55                       60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                       70                       75                       80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                       90                       95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                       105                       110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115                       120                       125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130                       135                       140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145                       150                       155                       160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
     165                       170                       175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180                       185                       190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195                       200                       205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
     210                       215                       220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     225                       230                       235                       240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     245                       250                       255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     260                       265                       270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     275                       280                       285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     290                       295                       300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     305                       310                       315                       320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     325                       330                       335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
     340                       345                       350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Asn Asp  
     355                       360                       365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     370                       375                       380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     385                       390                       395                       400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     405                       410                       415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     420                       425                       430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435                       440                       445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450                       455                       460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465                       470                       475                       480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485                       490                       495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val

-148-

	500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		
 <b>&lt;210&gt; 120</b>			
<b>&lt;211&gt; 312</b>			
<b>&lt;212&gt; PRT</b>			
<b>&lt;213&gt; Artificial Sequence</b>			
 <b>&lt;220&gt;</b>			
<b>&lt;223&gt; Mutant rep protein: rep40 350 GCT</b>			
 <b>&lt;400&gt; 120</b>			
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Leu Ala Arg Gly His Ser Leu			
305	310		

<b>&lt;210&gt; 121</b>			
<b>&lt;211&gt; 621</b>			
<b>&lt;212&gt; PRT</b>			
<b>&lt;213&gt; Artificial Sequence</b>			

-149-

<220>

<223> Mutant rep protein: rep78 354 GCC

<400> 121

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Ala Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
420 425 430  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
435 440 445  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln

-150-

450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 122

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 354 GCC

&lt;400&gt; 122

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		15
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		80
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Ala Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		160
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
		240

-151-

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

&lt;210&gt; 123

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 354 GCC

<400> 123  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala

-152-

Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
245						250							255		
260						265							270		
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
275						280							285		
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290						295						300			
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305						310					315			320	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
325						325					330			335	
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
340						340					345			350	
Phe	Ala	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
355						355					360			365	
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
370						370					375			380	
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385						385					390			395	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
405						405					410			415	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420						420					425			430	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435						435					440			445	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450						450					455			460	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465						465					470			475	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
485						485					490			495	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
500						500					505			510	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
515						515					520			525	
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
530						530					535				

&lt;210&gt; 124

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 354 GCC

&lt;400&gt; 124

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1						5					10			15	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							20			25			30		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						35					40			45	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
						50					55			60	
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
65						65					70			75	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
						85					90			95	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						100					105			110	

-153-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Ala Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

<210> 125

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 363 GCC

<400> 125

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn

-154-

195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Ala Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

&lt;210&gt; 126

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 52 363 GCC

-155-

<400> 126  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
   1               5               10               15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
   20              25              30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
   35              40              45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
   50              55              60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
   65              70              75              80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
   85              90              95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
   100             105             110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
   115             120             125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Ala Phe Pro Phe Asn Asp  
   130             135             140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
   145             150             155             160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
   165             170             175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
   180             185             190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
   195             200             205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
   210             215             220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
   225             230             235             240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
   245             250             255  
 Glu His Glu Phe Tyr Val Lys Gly Gly Ala Lys Lys Arg Pro Ala  
   260             265             270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
   275             280             285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
   290             295             300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
   305             310             315             320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
   325             330             335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
   340             345             350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
   355             360             365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
   370             375             380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
   385             390             395

&lt;210&gt; 127

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 363 GCC

&lt;400&gt; 127

-156-

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1					5				10						15
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
					20				25						30
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
					35				40						45
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
					50				55						60
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
					65				70						80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
					85				90						95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
					100				105						110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
					115				120						125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
					130				135						140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
					145				150						160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
					165				170						175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
					180				185						190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
					195				200						205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
					210				215						220
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
					225				230						240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
					245				250						255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					260				265						270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					275				280						285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
					290				295						300
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
					305				310						320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					325				330						335
Thr	Thr	Gly	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro	
					340				345						350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Ala	Phe	Pro	Phe	Asn	Asp
					355				360						365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370				375						380
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					385				390						400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					405				410						415
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					420				425						430
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					435				440						445
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
					450				455						460
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					465				470						475
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Ala	Lys	Lys	Arg	Pro	Ala	

-157-

485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500 505 510		
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515 520 525		
Arg Leu Ala Arg Gly His Ser Leu		
530 535		

<210> 128  
<211> 312  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 363 GCC

<400> 128	<400> 128	<400> 128
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1 5 10 15		
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20 25 30		
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35 40 45		
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50 55 60		
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65 70 75 80		
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
85 90 95		
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100 105 110		
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115 120 125		
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Ala Phe Pro Phe Asn Asp		
130 135 140		
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145 150 155 160		
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
- 165 170 175		
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180 185 190		
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195 200 205		
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210 215 220		
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225 230 235 240		
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
245 250 255		
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260 265 270		
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275 280 285		
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290 295 300		
Arg Leu Ala Arg Gly His Ser Leu		
305 310		

<210> 129  
<211> 621  
<212> PRT

-158-

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 364 GCT

<400> 129

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1					5				10						15
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
					20				25					30	
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
					35				40					45	
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
					50				55			60			
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70				75					80	
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
					85				90			95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
					100				105					110	
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
					115				120			125			
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
					130				135			140			
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145					150				155					160	
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
					165				170			175			
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
					180				185			190			
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
					195				200			205			
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
210					215				220						
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225					230				235			240			
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
					245				250			255			
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					260				265			270			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					275				280			285			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Glu	Leu	
290					295				300						
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305					310				315			320			
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					325				330			335			
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					340				345			350			
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Ala	Pro	Phe	Asn	Asp
					355				360			365			
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370				375			380			
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385					390				395			400			
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					405				410			415			
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					420				425			430			
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe

-159-

435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

<210> 130

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 364 GCT

<400> 130

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys	15
1	5
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala	30
20	25
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys	45
35	40
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln	60
50	55
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu	80
65	70
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala	95
85	90
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala	110
100	105
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro	125
115	120
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Ala Pro Phe Asn Asp	140
130	135
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala	160
145	150
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg	175
165	170
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val	190
180	185
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser	205
195	200
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe	220
210	215

-160-

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225           230           235           240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245           250           255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260           265           270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275           280           285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290           295           300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305           310           315           320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325           330           335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340           345           350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355           360           365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370           375           380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385           390           395

<210> 131

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 364 GCT

<400> 131

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1           5           10           15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20           25           30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35           40           45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50           55           60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65           70           75           80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85           90           95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100          105          110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115          120          125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130          135          140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145          150          155          160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165          170          175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180          185          190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195          200          205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210          215          220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys

-161-

225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr	Ile Ser Phe Asn Ala Ala		
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala	Leu Asp Asn Ala Gly Lys		
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp	Tyr Leu Val Gly Gln Gln		
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile	Tyr Lys Ile Leu Glu Leu		
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser	Val Phe Leu Gly Trp Ala		
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr	Ile Trp Leu Phe Gly Pro Ala		
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala	Ile Ala His Thr Val Pro		
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Ala	Pro Phe Asn Asp		
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly	Lys Met Thr Ala		
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg		
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln	Ile Asp Pro Thr Pro Val		
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile	Asp Gly Asn Ser		
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe		
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln		
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val		
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly	Gly Ala Lys Lys Arg Pro Ala		
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val	Arg Glu Ser Val		
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile	Asn Tyr Ala Asp		
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		

<210> 132

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 364 GCT

<400> 132

Met Glu Leu Val Gly Trp Leu Val Asp Lys	Gly Ile Thr Ser Glu Lys		
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser	Tyr Ile Ser Phe Asn Ala Ala		
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala	Leu Asp Asn Ala Gly Lys		
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp	Tyr Leu Val Gly Gln Gln		
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile	Tyr Lys Ile Leu Glu Leu		
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser	Val Phe Leu Gly Trp Ala		
85	90	95	

-162-

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Ala Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

<210> 133

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 367 GCC

<400> 133  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His

-163-

	180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Ala Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
565	570	575	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

&lt;210&gt; 134

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-164-

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 367 GCC

```

<400> 134
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
      5          10          15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
     20          25          30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
     35          40          45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
     50          55          60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
     65          70          75          80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
     85          90          95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
    100          105          110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
    115          120          125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Ala Asp
    130          135          140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
    145          150          155          160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
    165          170          175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
    180          185          190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
    195          200          205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
    210          215          220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
    225          230          235          240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
    245          250          255
Glu His Glu Phe Tyr Val Lys Gly Gly Ala Lys Lys Arg Pro Ala
    260          265          270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
    275          280          285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
    290          295          300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
    305          310          315          320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
    325          330          335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
    340          345          350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
    355          360          365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
    370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
    385          390          395

```

&lt;210&gt; 135

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

-165-

<223> Mutant rep protein: rep68 367 GCC

<400>	135		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
275	280	285	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
290	295	300	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Ala Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	

-166-

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
465 470 475 480  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
485 490 495  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
500 505 510  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
515 520 525  
Arg Leu Ala Arg Gly His Ser Leu  
530 535

<210> 136  
<211> 312  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep40 367 GCC

<400> 136  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Ala Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Leu Ala Arg Gly His Ser Leu  
305 310

-167-

<210> 137  
<211> 621  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Mutant rep protein: rep78 370 GCC  
  
<400> 137  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Ala Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415

-168-

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     420                        425                        430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435                        440                        445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450                        455                        460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465                        470                        475                        480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485                        490                        495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     500                        505                        510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     515                        520                        525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
     530                        535                        540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
     545                        550                        555                        560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
     565                        570                        575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
     580                        585                        590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
     595                        600                        605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
     610                        615                        620

<210> 138

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 370 GCC

<400> 138

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1                        5                        10                        15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20                        25                        30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                        40                        45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                        55                        60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                        70                        75                        80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                        90                        95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                       105                       110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115                       120                       125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130                       135                       140  
 Cys Ala Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145                       150                       155                       160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                       170                       175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                       185                       190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser

-169-

195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
305	310	315
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		320
325	330	335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
340	345	350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
355	360	365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
370	375	380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
385	390	395

&lt;210&gt; 139

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 370 GCC

&lt;400&gt; 139

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		15
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		80
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205

-170-

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Ala Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

&lt;210&gt; 140

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 370 GCC

&lt;400&gt; 140

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu

-171-

	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Ala Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Leu Ala Arg Gly His Ser Leu			
305	310		

<210> 141

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 376 GCG

<400> 141

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1				5					10					15	
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
				20				25					30		
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
				35				40				45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
				50			55			60					
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
				65		70			75				80		
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
				85			90					95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
				100			105					110			
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
				115			120				125				
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
				130		135			140						
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
	145				150				155				160		

-172-

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Ala Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460 465  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 580 585 590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 595 600 605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 610 615 620

<210> 142  
<211> 397

-173-

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep 52 376 GCG

<400> 142  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Ala Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
305 310 315 320  
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
325 330 335  
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
340 345 350  
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
355 360 365  
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
385 390 395

<210> 143  
<211> 536  
<212> PRT

-174-

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 376 GCG

<400> 143

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1					5				10					15	
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
					20			25					30		
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
					35			40				45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
					50			55			60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70				75			80			
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
					85				90			95			
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
					100			105			110				
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
					115			120			125				
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
					130			135			140				
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145					150				155			160			
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
					165				170			175			
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
					180			185			190				
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
					195			200			205				
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
					210			215			220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
					225			230			235			240	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
					245				250			255			
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					260			265			270				
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					275			280			285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
					290			295			300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305					310				315			320			
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					325				330			335			
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					340			345			350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					355			360			365				
Cys	Val	Asp	Lys	Met	Val	Ile	Ala	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370			375			380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385					390				395			400			
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					405				410			415			
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
					420			425			430				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe

-175-

435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 144

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep 40 376 GCG

<400> 144

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Ala Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	

-176-

Arg Leu Ala Arg Gly His Ser Leu  
305 310

<210> 145  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep 78 381 GCG

<400> 145  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Ala Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg

-177-

385	390	395	400
Val Asp Gln Lys Cys	Lys Ser Ser Ala Gln	Ile Asp Pro Thr Pro	Val
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met	Cys Ala Val Ile Asp Gly	Asn Ser	
420	425	430	
Thr Thr Phe Glu His Gln Gln	Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His	Asp Phe Gly Lys Val	Thr Lys Gln	
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys	Asp His Val Val	Glu Val	
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly	Gly Ala Lys Lys Arg	Pro Ala	
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu	Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala	Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His	Val Gly Met Asn	Leu Met Leu	
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn	Gln Asn Ser Asn	Ile Cys	
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu	Glu Cys Phe Pro Val	Ser Glu	
565	570	575	
Ser Gln Pro Val Ser Val Val Lys	Lys Ala Tyr Gln Lys	Leu Cys Tyr	
580	585	590	
Ile His His Ile Met Gly Lys Val	Pro Asp Ala Cys	Thr Ala Cys Asp	
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys	Ile Phe Glu Gln		
610	615	620	

&lt;210&gt; 146

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 52 381 GCG

&lt;400&gt; 146

Met Glu Leu Val Gly Trp Leu Val Asp Lys	Gly Ile Thr Ser Glu Lys		
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser	Tyr Ile Ser Phe Asn Ala Ala		
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys	Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp	Tyr Leu Val Gly Gln Gln		
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg	Ile Tyr Lys Ile Leu Glu Leu		
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala	Ser Val Phe Leu Gly Trp Ala		
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn	Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu	Ala Ile Ala His Thr Val Pro		
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn	Glu Asn Phe Pro Phe Asn Asp		
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp	Glu Glu Gly Ala Met Thr Ala		
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile	Leu Gly Gly Ser Lys Val Arg		
165	170	175	

-178-

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

&lt;210&gt; 147

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 68 381 GCG

<400> 147  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His

-179-

180	185	190
Leu Thr His Val Ser Gln Thr Gln	Glu Gln Asn Lys Glu Asn Gln Asn	
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Ala Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 148

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 381 GCG

<400> 148

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20	25	30

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45

-180-

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Ala Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

<210> 149

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 382 GCG

<400> 149

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly

-181-

130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Ala Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		560
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

-182-

<210> 150  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep 52 382 GCG

<400> 150  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Ala Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
305 310 315 320  
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
325 330 335  
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
340 345 350  
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
355 360 365  
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
385 390 395

<210> 151

-183-

<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep 68 382 GCG

<400> 151  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Ala Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400  
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
405 410 415  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser

-184-

420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 152

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 382 GCG

400	405	410
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Ala Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Ser Lys Val Arg		
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285

-185-

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Leu Ala Arg Gly His Ser Leu  
305 310

<210> 153  
<211> 621  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep78 389 GCG

<400> 153  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala

-186-

370	375	380
Lys Val Val Glu Ala Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln	Ile Asp Pro Thr Pro Val	400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		480
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		495
485	490	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		575
565	570	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

<210> 154  
<211> 397  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep52 389 GCG

<400> 154		
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		15
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		80
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
		160

-187-

Lys Val Val Glu Ala Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
           165                 170                 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
           180                 185                 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
           195                 200                 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
           210                 215                 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
           225                 230                 235                 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
           245                 250                 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
           260                 265                 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
           275                 280                 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
           290                 295                 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
           305                 310                 315                 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
           325                 330                 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
           340                 345                 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
           355                 360                 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
           370                 375                 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
           385                 390                 395

&lt;210&gt; 155

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep68 389 GCG

<400> 155  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1              5                 10                 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20              25                 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35              40                 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50              55                 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65              70                 75                 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85              90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100            105                 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115            120                 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130            135                 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145            150                 155                 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu

-188-

Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
165							170							175	
180							185							190	
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
195							200							205	
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
210							215							220	
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225							230							235	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
245							250							255	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
260							265							270	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
275							280							285	
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290							295							300	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305							310							315	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
325							325							330	
Thr	Thr	Gly	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro	
340							345							350	
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
355							360							365	
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
370							375							380	
Lys	Val	Val	Glu	Ala	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385							390							395	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
405							405							410	
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
420							420							425	
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
435							435							440	
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
450							450							455	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
465							465							470	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
485							485							490	
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
500							500							505	
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
515							515							520	
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu							525	
530							530							535	

&lt;210&gt; 156

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 389 GCG

&lt;400&gt; 156

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys

1 5 10 15

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala

20 25 30

-189-

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                  40                  45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                  55                  60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                  70                  75                  80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                  90                  95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                 105                 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115                 120                 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130                 135                 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145                 150                 155                 160  
 Lys Val Val Glu Ala Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165                 170                 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180                 185                 190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     195                 200                 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210                 215                 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225                 230                 235                 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245                 250                 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260                 265                 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275                 280                 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290                 295                 300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305                 310

&lt;210&gt; 157

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 407 GCC

&lt;400&gt; 157

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1                  5                  10                  15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20                 25                 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                 40                 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                 55                 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                 70                 75                 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                 90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100                105                110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu

-190-

115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala	Gly Gly Gly	
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu	Leu Pro Lys	
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu	Gln Tyr Leu	
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val	Ala Gln His	
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu	Asn Gln Asn	
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser	Ala Arg Tyr	
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr	Ser Glu Lys	
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe	Asn Ala Ala	
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn	Ala Gly Lys	
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val	Gly Gln Gln	
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile	Leu Glu Leu	
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu	Gly Trp Ala	
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe	Gly Pro Ala	
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His	Thr Val Pro	
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro	Phe Asn Asp	
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys	Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser	Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ala Ser Ala Gln Ile Asp Pro	Thr Pro Val	
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met	Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val	Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val	Val Glu Val	
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys	Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn	Leu Met Leu	
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser	Asn Ile Cys	
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro	Val Ser Glu	
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys	Leu Cys Tyr	
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr	Ala Cys Asp	
595	600	605

-191-

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
610 615 620

<210> 158

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 407 GCC

<400> 158

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
1 5 10 15  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
20 25 30  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
35 40 45  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
50 55 60  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
65 70 75 80  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
85 90 95  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
100 105 110  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
115 120 125  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
130 135 140  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
145 150 155 160  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
165 170 175  
Val Asp Gln Lys Cys Lys Ala Ser Ala Gln Ile Asp Pro Thr Pro Val  
180 185 190  
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
195 200 205  
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
210 215 220  
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
225 230 235 240  
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
245 250 255  
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
260 265 270  
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
275 280 285  
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
290 295 300  
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
305 310 315 320  
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
325 330 335  
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
340 345 350  
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
355 360 365  
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln

-192-

385	390	395
-----	-----	-----

<210> 159  
<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein: rep68 407 GCC

<400> 159  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
370 375 380  
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
385 390 395 400

-193-

<210> 160

<211> 312

<212> PRT

### <213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 407 GCC

<400> 160  
Met. Glu. I

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1				5					10					15	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
						20		25					30		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						35		40				45			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
						50		55			60				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
						65	-	70		75			80		
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
						85			90			95			
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						100		105				110			
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
						115		120			125				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
						130		135			140				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
						145		150		155			160		
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
						165			170			175			
Val	Asp	Gln	Lys	Cys	Lys	Ala	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
						180		185			190				
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
						195		200			205				
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
						210		215			220				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
						225		230		235			240		
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
						245			250			255			
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala

-194-

260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Leu Ala Arg Gly His Ser Leu		
305	310	
<210> 161		
<211> 621		
<212> PRT		
<213> Artificial Sequence		
<220>		
<223> Mutant rep protein: rep78 411 GCA		
<400> 161		
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp		
1	5	10
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu		
20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile		
35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu		
50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350

-195-

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     355                         360                         365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     370                         375                         380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     385                         390                         395                         400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ala Asp Pro Thr Pro Val  
     405                         410                         415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
     420                         425                         430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435                         440                         445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450                         455                         460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465                         470                         475                         480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485                         490                         495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     500                         505                         510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     515                         520                         525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
     530                         535                         540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
     545                         550                         555                         560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
     565                         570                         575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
     580                         585                         590  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
     595                         600                         605  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
     610                         615                         620

&lt;210&gt; 162

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: 52 411 GCA

<400> 162  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1                         5                                 10                         15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20                         25                                 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35                         40                                 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50                         55                                 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65                         70                                 75                         80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85                         90                                 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100                         105                                 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115                         120                                 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp

-196-

130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		160
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ala Asp Pro Thr Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
305	310	315
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		320
325	330	335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
340	345	350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
355	360	365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
370	375	380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
385	390	395

<210> 163

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: 68 411 GCA

<400> 163

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly			
130	135	140	

-197-

Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145				150					155					160	
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
				165					170					175	
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
				180				185					190		
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
				195				200				205			
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
				210			215				220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
				225			230			235				240	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
				245			250						255		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
				260			265					270			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
				275			280				285				
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
				290			295				300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
				305			310			315				320	
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
				325			330				335				
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ile	Ala	His	Thr	Val	Pro	
				340			345				350				
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
				355			360				365				
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
				370			375				380				
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
				385			390			395				400	
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ala	Asp	Pro	Thr	Pro	Val
				405			410						415		
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
				420			425					430			
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
				435			440				445				
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
				450			455			460					
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
				465			470			475				480	
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
				485			490					495			
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
				500			505				510				
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
				515			520				525				
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
				530			535								

&lt;210&gt; 164

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep 40 411 GCA

&lt;400&gt; 164

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys

-198-

1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ala Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Leu Ala Arg Gly His Ser Leu			
305	310		

```
<210> 165  
<211> 621  
<212> PRT  
<213> Artificial Sequence
```

<220>  
<223> Mutant rep protein: rep78 414 GCT

```

<400> 165
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
      1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
      20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
      35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
      50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
      65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
      85          90          95

```

-199-

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 325 330 335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Ala Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 530 535 540  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 545 550 555 560  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 565 570 575  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr

-200-

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp 595 600 605 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln 610 615 620	580                    585                    590
 <b>&lt;210&gt; 166</b>	
<b>&lt;211&gt; 397</b>	
<b>&lt;212&gt; PRT</b>	
<b>&lt;213&gt; Artificial Sequence</b>	
 <b>&lt;220&gt;</b>	
<b>&lt;223&gt; Mutant rep protein: rep52 414 GCT</b>	
 <b>&lt;400&gt; 166</b>	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys 1                5                10                15	
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala 20                25                30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys 35                40                45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln 50                55                60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu 65                70                75                80	
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala 85                90                95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala 100              105              110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro 115              120              125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp 130              135              140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 145              150              155              160	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 165              170              175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Ala Pro Val 180              185              190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 195              200              205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe 210              215              220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln 225              230              235              240	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val 245              250              255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala 260              265              270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val 275              280              285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp 290              295              300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu 305              310              315              320	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys 325              330              335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu 340              345              350	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr 355              360              365	

-201-

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
385 390 395

<210> 167  
<211> 536  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Mutant rep protein:rep68 414 GCT

<400> 167  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
305 310 315 320  
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
325 330 335  
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
340 345 350  
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
355 360 365  
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala

-202-

370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Ala Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 168

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 414 GCT

&lt;400&gt; 168

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		15
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		80
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		160
165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Ala Pro Val		
180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
		240

-203-

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305 310

&lt;210&gt; 169

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: 78 420 GCT

&lt;400&gt; 169

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115 120 125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
 130 135 140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145 150 155 160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165 170 175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180 185 190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195 200 205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210 215 220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225 230 235 240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245 250 255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260 265 270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275 280 285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290 295 300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 305 310 315 320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala

-204-

325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His	Thr Val Pro	
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro	Phe Asn Asp	
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly	Lys Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly	Gly Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro	Thr Pro Val	
405	410	415
Ile Val Thr Ala Asn Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val	
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly	Gly Ala Lys Lys Arg Pro Ala	
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly	Met Asn Leu Met Leu	
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser	Asn Ile Cys	
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro	Val Ser Glu	
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys	Leu Cys Tyr	
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr	Ala Cys Asp	
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

<210> 170

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep 52 420 GCT

<400> 170

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
1	5	10
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
20	25	30
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
35	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
50	55	60
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
65	70	75
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
100	105	110

-205-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ala Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

<210> 171

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 420 GCT

<400> 171

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100 105 110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu

-206-

115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ala Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

&lt;210&gt; 172

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 420 GCT

-207-

<400> 172  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
   1               5               10               15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
   20               25               30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
   35               40               45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
   50               55               60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
   65               70               75               80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
   85               90               95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
   100              105              110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
   115              120              125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
   130              135              140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
   145              150              155              160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
   165              170              175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
   180              185              190  
 Ile Val Thr Ala Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
   195              200              205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
   210              215              220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
   225              230              235              240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
   245              250              255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
   260              265              270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
   275              280              285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
   290              295              300  
 Arg Leu Ala Arg Gly His Ser Leu  
   305              310

&lt;210&gt; 173

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 421 GCC

&lt;400&gt; 173

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
   1               5               10               15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
   20               25               30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
   35               40               45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
   50               55               60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
   65               70               75               80

-208-

Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
85								90							95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
		100						105							110
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
	115						120							125	
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
	130					135					140				
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145						150				155					160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
	165							170						175	
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
	180						185							190	
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
	195						200					205			
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
	210					215					220				
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225						230				235					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
	245							250					255		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
	260						265					270			
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
	275						280					285			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290						295					300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305						310				315			320		
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
	325							330					335		
Thr	Thr	Gly	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro	
	340						345					350			
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
	355						360					365			
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
	370					375				380					
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
385						390			395				400		
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
	405							410					415		
Ile	Val	Thr	Ser	Ala	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
	420						425					430			
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
	435						440					445			
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
	450					455				460					
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
	465					470			475				480		
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
	485							490					495		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
	500						505					510			
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
	515						520					525			
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
	530					535				540					
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
	545					550				555				560	
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu

-209-

Ser Gln Pro Val Ser Val Val Lys	565	570	575
580	585	590	
Ile His His Ile Met Gly Lys Val		Pro Asp Ala Cys Thr	Ala Cys Asp
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp		Cys Ile Phe Glu Gln	
610	615.	620	
<210> 174			
<211> 397			
<212> PRT			
<213> Artificial Sequence			
<220>			
<223> Mutant rep protein: rep52 421 GCC			
<400> 174			
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Ala Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	310	315	320
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	330	335	
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	345	350	

-210-

```

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
355          360          365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
370          375          380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385          390          395

<210> 175
<211> 536
<212> PRT
<213> Artificial Sequence

<220>
<223> Mutant rep protein: rep68 421 GCC

<400> 175
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1           5           10          15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20          25          30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35          40          45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50          55          60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65          70          75          80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85          90          95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100         105         110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115         120         125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130         135         140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145         150         155         160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165         170         175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180         185         190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195         200         205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210         215         220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225         230         235         240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245         250         255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260         265         270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275         280         285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290         295         300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305         310         315         320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325         330         335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340         345         350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp

```

-211-

355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly	Lys Met Thr Ala	
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly	Ser Lys Val Arg	
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro	Thr Pro Val	400
405	410	415
Ile Val Thr Ser Ala Thr Asn Met Cys Ala Val Ile Asp	Gly Asn Ser	
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg	Met Phe Lys Phe	
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly	Lys Val Thr Lys Gln	
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His	Val Val Glu Val	465
470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys	Lys Arg Pro Ala	485
490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg	Glu Ser Val	
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn	Tyr Ala Asp	
515	520	525
Arg Leu Ala Arg Gly His Ser Leu		
530	535	

<210> 176

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 421 GCC

<400> 176

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	-	40	45
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	-	90	95
Thr Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	-	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Ala Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	

-212-

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225                   230                   235                   240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245                   250                   255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260                   265                   270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275                   280                   285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290                   295                   300  
 Arg Leu Ala Arg Gly His Ser Leu  
 305                   310

<210> 177  
 <211> 621  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep78 422 GCC

<400> 177  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1               5                   10                   15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20               25                   30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35               40                   45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50               55                   60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65               70                   75                   80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85               90                   95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
 100              105                   110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
 115              120                   125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
 130              135                   140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
 145              150                   155                   160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
 165              170                   175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
 180              185                   190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
 195              200                   205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
 210              215                   220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 225              230                   235                   240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 245              250                   255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 260              265                   270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 275              280                   285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 290              295                   300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala

-213-

305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr-Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Ala Asn Met Cys Ala Val Ile Asp Gly Asn Ser			
420	425	430	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
565	570	575	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

&lt;210&gt; 178

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 422 GCC

&lt;400&gt; 178

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	

-214-

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Ala Asn Met Cys Ala Val Ile Asp Gly Asn Ser  
 195 200 205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 210 215 220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 225 230 235 240  
 Glu Val Lys Asp Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 245 250 255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 260 265 270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 275 280 285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 290 295 300  
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
 305 310 315 320  
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
 325 330 335  
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
 340 345 350  
 Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
 355 360 365  
 Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
 370 375 380  
 Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
 385 390 395

<210> 179

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 422 GCC

<400> 179

Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
 1 5 10 15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
 20 25 30  
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
 35 40 45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
 50 55 60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
 65 70 75 80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
 85 90 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile

-215-

Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
100							105								110
115							120								125
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
130							135								140
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145							150								155
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
							165								170
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
							180								185
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
							195								200
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
							210								215
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
							225								230
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
							245								250
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
							260								265
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
							275								280
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
							290								295
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
							305								310
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
							325								330
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
							340								345
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
							355								360
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
							370								375
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Ser	Lys	Val	Arg	
							385								390
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
							405								410
Ile	Val	Thr	Ser	Asn	Ala	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
							420								425
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
							435								440
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
							450								455
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
							465								470
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
							485								490
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
							500								505
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
							515								520
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
							530								535

&lt;210&gt; 180

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

-216-

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 422 GCC

&lt;400&gt; 180

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1					5				10					15	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
						20			25					30	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
						35			40					45	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Gly	Gln	Gln	
						50			55			60			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
						65			70			75		80	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
						85			90			95			
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						100			105			110			
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
						115			120			125			
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
						130			135			140			
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
						145			150			155		160	
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
						165			170			175			
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
						180			185			190			
Ile	Val	Thr	Ser	Asn	Ala	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
						195			200			205			
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
						210			215			220			
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
						225			230			235		240	
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
						245			250			255			
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
						260			265			270			
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
						275			280			285			
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
						290			295			300			
Arg	Leu	Ala	Arg	Gly	His	Ser	Leu								
						305			310						

&lt;210&gt; 181

&lt;211&gt; 621

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep78 424 GCG

&lt;400&gt; 181

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1					5				10				15		
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
						20			25			30			
Lys	Glu	Trp	Glu	Leu	Phe	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Ile	
						35			40			45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu

-217-

50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val		
65	70	75
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu		80
85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile		
100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Thr Asn Met Glu Gln Tyr Leu		160
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		240
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		320
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		400
405	410	415
Ile Val Thr Ser Asn Thr Asn Ala Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		480
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540

-218-

Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
545				550					555						560
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
				565				570							575
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
				580			585				590				
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
				595			600				605				
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
				610			615				620				

&lt;210&gt; 182

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<220>  
<223> Mutant rep protein: rep52 424 GCG

<400>	182														
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1				5					10						15
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
				20				25						30	
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
				35			40							45	
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
				50			55							60	
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
65				70			75							80	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
	85			90			95								
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
	100			105			110								
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
	115			120			125								
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
	130		135		140										
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
145		150		155		160									
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
	165		170		175										
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
	180		185		190										
Ile	Val	Thr	Ser	Asn	Thr	Asn	Ala	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser
	195		200		205										
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
	210		215		220										
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
225		230		235		240									
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
	245		250		255										
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
	260		265		270										
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
	275		280		285										
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
	290		295		300										
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
	305		310		315									320	
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys

-219-

Phe	Thr	His	Gly	325	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
				340					345						350	
Ser	Gln	Pro	Val	355	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
				370					360						365	
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp	
									375						380	
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln				
									390						395	

<210> 183

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 424 GCG

<400> 183

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1				5				10			15				
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
					20			25			30				
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
					35			40			45				
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
					50			55			60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70			75			80				
Gln	Phe	Glu	Lys	Gly	Ser	Tyr	Phe	His	Met	His	Val	Leu	Glu		
	85						90						95		
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
	100						105						110		
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
	115						120						125		
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	
	130						135						140		
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
	145						150						155		160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
	165						170						175		
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
	180						185						190		
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
	195						200						205		
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
	210						215						220		
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
	225						230						235		240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
	245						250						255		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
	260						265						270		
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
	275						280						285		
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
	290						295						300		
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
	305						310						315		320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
	325												330		335

-220-

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 340 345 350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 355 360 365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 370 375 380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 385 390 395 400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 405 410 415  
 Ile Val Thr Ser Asn Thr Asn Ala Cys Ala Val Ile Asp Gly Asn Ser  
 420 425 430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
 435 440 445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
 450 455 460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
 465 470 475 480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
 485 490 495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
 500 505 510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
 515 520 525  
 Arg Leu Ala Arg Gly His Ser Leu  
 530 535

<210> 184

<211> 312

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep40 424 GCG

<400> 184  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
 1 5 10 15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
 20 25 30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
 35 40 45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
 50 55 60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 65 70 75 80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
 85 90 95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
 100 105 110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
 115 120 125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 130 135 140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 145 150 155 160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
 165 170 175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
 180 185 190  
 Ile Val Thr Ser Asn Thr Asn Ala Cys Ala Val Ile Asp Gly Asn Ser

-221-

195	200	205
Thr Thr Phe Glu His Gln Gln	Pro Leu Gln Asp Arg Met Phe Lys Phe	
210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		240
245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
290	295	300
Arg Leu Ala Arg Gly His Ser Leu		
305	310	

<210> 185

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep78 428 GCT

<400> 185			
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
20	25	30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
35	40	45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
50	55	60	
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
85	90	95	
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			
100	105	110	
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu			
115	120	125	
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly			
130	135	140	
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys			
145	150	155	160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu			
165	170	175	
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His			
180	185	190	
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn			
195	200	205	
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr			
210	215	220	
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
225	230	235	240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
245	250	255	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
260	265	270	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
275	280	285	

-222-

Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
290					295						300				
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305					310					315					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
					325				330					335	
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					340				345				350		
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					355				360				365		
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala
					370				375			380			
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg
					385				390			395			400
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val
					405				410				415		
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ala	Asp	Gly	Asn	Ser
					420				425				430		
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe
					435				440				445		
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln
					450				455			460			
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val
					465				470			475			480
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala
					485				490				495		
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val
					500				505				510		
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp
					515				520				525		
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu
					530				535			540			
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys
					545				550			555			560
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu
					565				570				575		
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr
					580				585			590			
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp
					595				600			605			
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln			
					610				615			620			

&lt;210&gt; 186

&lt;211&gt; 397

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep52 428 GCT

&lt;400&gt; 186

Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
1					5				10					15	
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
					20				25				30		
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					35				40				45		
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					50				55			60			
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu

-223-

65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	85	90	95
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	100	105	110
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	115	120	125
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	130	135	140
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	145	150	155
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	165	170	175
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	180	185	190
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ala Asp Gly Asn Ser			
195	195	200	205
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	210	215	220
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	225	230	235
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	245	250	255
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	260	265	270
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	275	280	285
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	290	295	300
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
305	305	310	315
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
325	325	330	335
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
340	340	345	350
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
355	355	360	365
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
370	370	375	380
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
385	385	390	395

<210> 187

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 428 GCT

<400> 187

Thr	Ala	Gly	Phe	Tyr	Glu	Ile	Val	Ile	Lys	Val	Pro	Ser	Asp	Leu	Asp
1				5					10					15	
Glu	His	Leu	Pro	Gly	Ile	Ser	Asp	Ser	Phe	Val	Asn	Trp	Val	Ala	Glu
				20				25					30		
Lys	Glu	Trp	Glu	Leu	Pro	Pro	Asp	Ser	Asp	Met	Asp	Leu	Asn	Leu	Ile
				35				40				45			
Glu	Gln	Ala	Pro	Leu	Thr	Val	Ala	Glu	Lys	Leu	Gln	Arg	Asp	Phe	Leu
				50				55			60				
Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70				75					80	

-224-

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
       85                 90                 95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
       100                105                110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
       115                120                125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
       130                135                140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
       145                150                155                160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
       165                170                175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
       180                185                190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
       195                200                205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
       210                215                220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
       225                230                235                240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
       245                250                255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
       260                265                270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
       275                280                285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
       290                295                300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
       305                310                315                320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
       325                330                335  
 Thr Thr Gly Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
       340                345                350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
       355                360                365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
       370                375                380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
       385                390                395                400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
       405                410                415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ala Asp Gly Asn Ser  
       420                425                430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
       435                440                445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
       450                455                460  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
       465                470                475                480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
       485                490                495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
       500                505                510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
       515                520                525  
 Arg Leu Ala Arg Gly His Ser Leu  
       530                535

<210> 188  
 <211> 312

-225-

<212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep40 428 GCT

<400> 188  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     1              5             10             15  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     20             25             30  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     35             40             45  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     50             55             60  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     65             70             75             80  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     85             90             95  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     100            105            110  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     115            120            125  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     130            135            140  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     145            150            155            160  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     165            170            175  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     180            185            190  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ala Asp Gly Asn Ser  
     195            200            205  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     210            215            220  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     225            230            235            240  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     245            250            255  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     260            265            270  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     275            280            285  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp  
     290            295            300  
 Arg Leu Ala Arg Gly His Ser Leu  
     305            310

<210> 189  
 <211> 621  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Mutant rep protein: rep78 429 GCC

<400> 189  
 Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
     1              5             10             15  
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
     20             25            30

-226-

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
     35                40                45  
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
     50                55                60  
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
     65                70                75                80  
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
     85                90                95  
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
     100               105               110  
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
     115               120               125  
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly  
     130               135               140  
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
     145               150               155               160  
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
     165               170               175  
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
     180               185               190  
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
     195               200               205  
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
     210               215               220  
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
     225               230               235               240  
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
     245               250               255  
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
     260               265               270  
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
     275               280               285  
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
     290               295               300  
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala  
     305               310               315               320  
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala  
     325               330               335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro  
     340               345               350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
     355               360               365  
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
     370               375               380  
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg  
     385               390               395               400  
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val  
     405               410               415  
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Ala Gly Asn Ser  
     420               425               430  
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe  
     435               440               445  
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln  
     450               455               460               480  
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val  
     465               470               475               480  
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala  
     485               490               495  
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val  
     500               505               510  
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp

-227-

515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn	Leu Met Leu	
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser	Asn Ile Cys	
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro	Val Ser Glu	
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys	Leu Cys Tyr	
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr	Ala Cys Asp	
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620

<210> 190

<211> 397

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep52 429 GCC

<400> 190

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
180	185	190	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Ala Gly Asn Ser			
195	200	205	
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
210	215	220	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
225	230	235	240
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
245	250	255	
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
260	265	270	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
275	280	285	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
290	295	300	

-228-

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu  
305 310 315 320  
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys  
325 330 335  
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu  
340 345 350  
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr  
355 360 365  
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp  
370 375 380  
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln  
385 390 395

<210> 191

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant rep protein: rep68 429 GCC

<400> 191  
Thr Ala Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp  
1 5 10 15  
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu  
20 25 30  
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile  
35 40 45  
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu  
50 55 60  
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val  
65 70 75 80  
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu  
85 90 95  
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile  
100 105 110  
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu  
115 120 125  
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly  
130 135 140  
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys  
145 150 155 160  
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu  
165 170 175  
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His  
180 185 190  
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn  
195 200 205  
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr  
210 215 220  
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys  
225 230 235 240  
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala  
245 250 255  
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys  
260 265 270  
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln  
275 280 285  
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
290 295 300  
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala

-229-

305	310	315	320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
325	330	335	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
340	345	350	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
355	360	365	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
370	375	380	
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
385	390	395	400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val			
405	410	415	
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Ala Gly Asn Ser			
420	425	430	
Thr Thr Phe-Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe			
435	440	445	
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln			
450	455	460	
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val			
465	470	475	480
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala			
485	490	495	
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val			
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Leu Ala Arg Gly His Ser Leu			
530	535		

&lt;210&gt; 192

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Mutant rep protein: rep40 429 GCC

&lt;400&gt; 192

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys			
1	5	10	15
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala			
20	25	30	
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys			
35	40	45	
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln			
50	55	60	
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu			
65	70	75	80
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala			
85	90	95	
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala			
100	105	110	
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro			
115	120	125	
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp			
130	135	140	
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala			
145	150	155	160
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg			
165	170	175	